EXHIBIT A

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

ZTE Corporation, ZTE (USA), Inc., and ZTE (TX), Inc.,

Petitioner,

v.

WSOU Investments LLC D/B/A Brazos Licensing and Development,

Patent Owner.

U.S. Patent No. 9,294,060
PCT Filing Date: May 25, 2010
Issue Date: Mar. 22, 2016

Case No. IPR2021-00697

PETITION FOR *INTER PARTES* REVIEW
OF U.S. PATENT NO. 9,294,060

TABLE OF CONTENTS

LIST OF EXHIBITS

Exhibit	Description
Ex-1001	U.S. Patent No. 9,294,060 (the ’060 patent)
Ex-1002	Prosecution File History of U.S. Patent No. 9,294,060
Ex-1003	Declaration of Nir Regev
Ex-1004	Curriculum Vitae of Nir Regev
Ex-1005	U.S. Patent No. 7,359,854 (“*Nilsson*”)
Ex-1006	U.S. Patent No. 8,160,889 (“*Iser*”)
Ex-1007	U.S. Patent App. Pub. No. 2006/0282262 (“*Vos*”)
Ex-1008	J. *Kontio*, L. Laaksonen and P. Alku, "Neural Network-Based Artificial Bandwidth Expansion of Speech," in *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15, no. 3, pp. 873-881, March 2007, doi: 10.1109/TASL.2006.885934. (“*Kontio*”)
Ex-1009	ETSI Standard, ETSI ES 201 108 V1.1.3 (2003-09), “Speech Processing, Transmission and Quality Aspects (STQ); Distributed speech recognition; Front-end feature extraction algorithm; Compression algorithms”, (“*ETSI201.108*”)
Ex-1010	Screen shot of ETSI (European Telecommunications Standards Institute) Work Programme showing the publication date of the *ETSI-201.108*
Ex-1011	U.S. Patent No. 6,711,536 B2
Ex-1012	S. Umesh, L. Cohen and D. Nelson, "Frequency warping and the Mel scale," in *IEEE Signal Processing Letters*, vol. 9, no. 3, pp. 104-107, March 2002, doi: 10.1109/97.995829.
Ex-1013	A. M. De Lima Araujo and F. Violaro, "Formant frequency estimation using a Mel-scale LPC algorithm," *ITS'98 Proceedings. SBT/IEEE International Telecommunications Symposium* (Cat. No.98EX202), Sao Paulo, Brazil, 1998, pp. 207-212 vol.1, doi: 10.1109/ITS.1998.713118.
Ex-1014	E. Karlsson and M. Hayes, "Least squares ARMA modeling of linear time-varying systems: Lattice filter structures and fast RLS algorithms," in *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, no. 7, pp. 994-1014, July 1987, doi: 10.1109/TASSP.1987.1165246.
Ex-1015	R. Badeau and B. David, "Weighted maximum likelihood autoregressive and moving average spectrum modeling," 2008 *IEEE International Conference on Acoustics, Speech and Signal*

	Processing, Las Vegas, NV, USA, 2008, pp. 3761-3764, doi: 10.1109/ICASSP.2008.4518471.
Ex-1016	Bernd *Iser*, Gerhard Schmidt, "Neural Networks Versus Codebooks in an Application for Bandwidth Extension of Speech Signals,", Neural Networks Versus Codebooks in an Application for Bandwidth Extension of Speech Signals ("*Iser-2*")
Ex-1017	Kornagel, Ulrich, "Techniques for artificial bandwidth extension of telephone speech", *Signal Processing* 86.6, pp. 1296-1306 (2006) ("*Kornagel*")
Ex-1018	*Iser* Bernd, Gerhard Schmidt, and Wolfgang Minker, Bandwidth extension of speech signals. Vol. 13, Springer Science & Business Media, page 15, and pages 77-94, 2008 ("*Iser-3*")

I. INTRODUCTION

ZTE Corporation, ZTE (USA), Inc., and ZTE (TX), Inc., (collectively "ZTE") requests *inter partes* review of claims 1-18 of U.S. Patent No. 9,294,060 (Ex-1001), assigned to WSOU Investments LLC ("WSOU")[1].

The '060 patent describes a method for extending the bandwidth of an audio signal, but the disclosed methods for bandwidth extension were known in the prior art.

A method comprising:

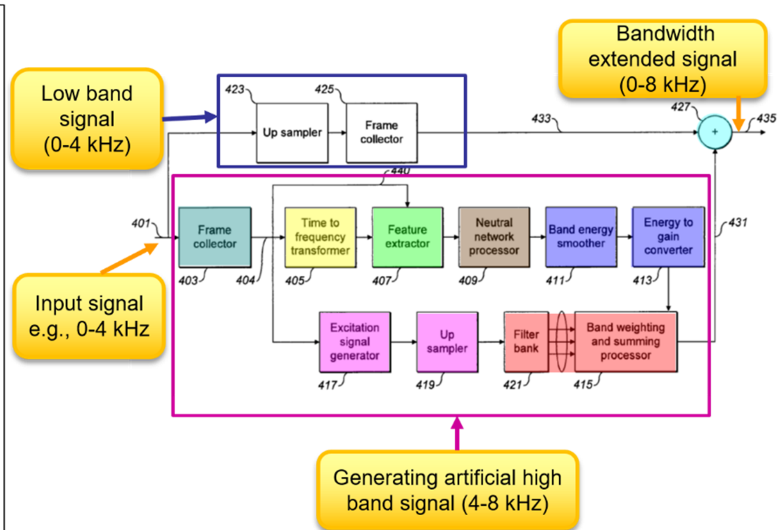
generating an excitation signal from an audio signal, wherein in the audio signal comprises a plurality of frequency components;

extracting a feature vector from the audio signal, wherein the feature vector comprises at least one frequency domain component feature and at least one time domain component feature;

determining at least one spectral shape parameter from the feature vector, wherein the at least one spectral shape parameter corresponds to a sub band signal comprising frequency components which belong to a further plurality of frequency components; and

generating the sub band signal by filtering the excitation signal through a filter bank and weighting the filtered excitation signal with the at least one spectral shape parameter,

wherein the spectral shape parameter is a sub band energy level value and the sub band energy level value is attenuated when the power of the audio signal approaches an estimate of the level of noise in the audio signal.

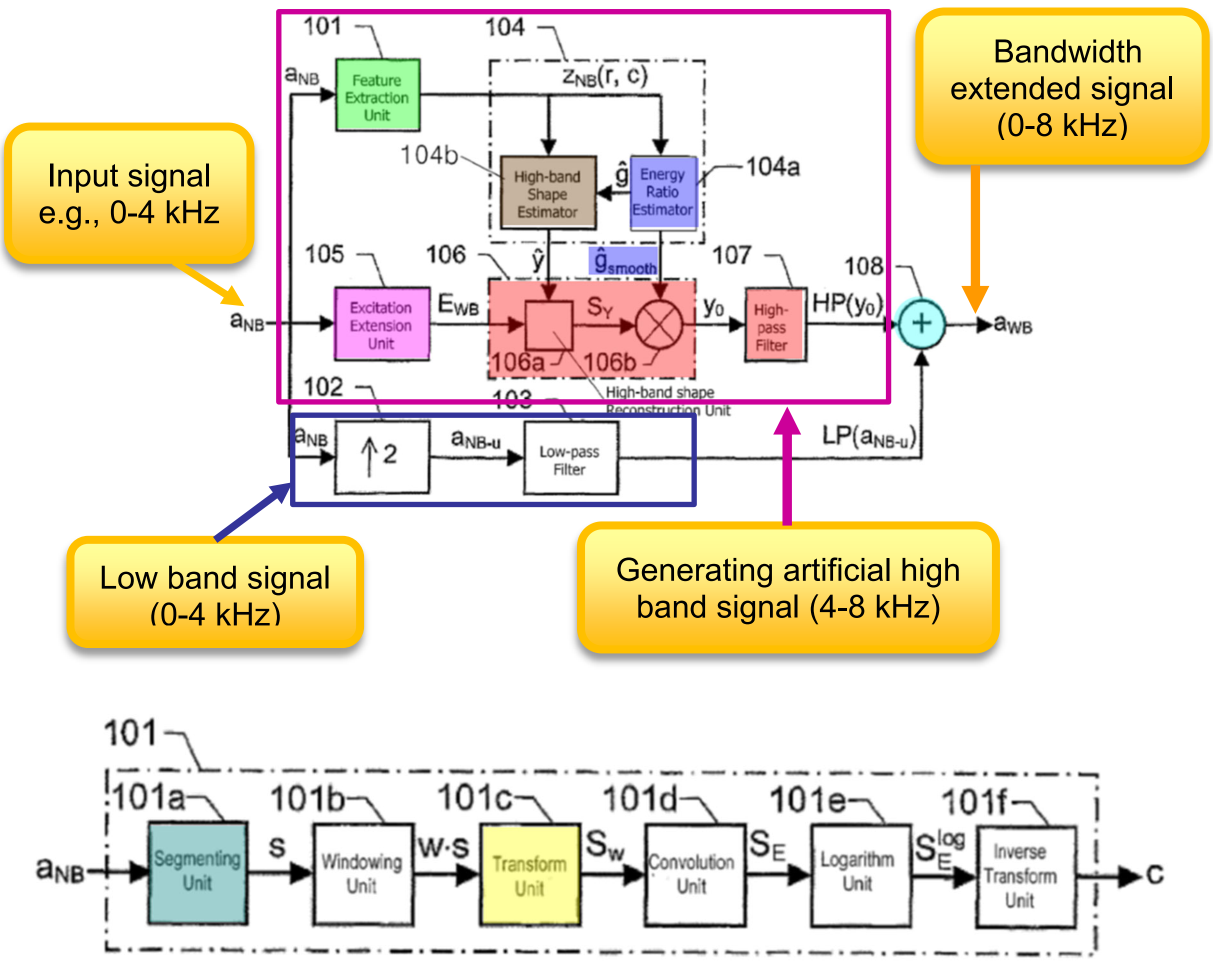


'060 Patent, Claim 1 and Figure 4 (Annotated)

Nilsson (Ex-1005) and *Vos* (Ex-1007) disclose methods for extending the bandwidth of an audio signal like those claimed in '060 patent. Alone or when combined in routine, predictable ways, *Nilsson*, *Vos*, *Iser*, *Kontio*, and

[1] WSOU asserts the '060 patent against ZTE in WSOU Investments LLC v. ZTE Corporation et al., 6:20-cv-00493-ADA (W.D. Tex.) ("Litigation").

ETSI201.108 —none considered by the USPTO for ’060 patent—render obvious all features of the challenged claims.

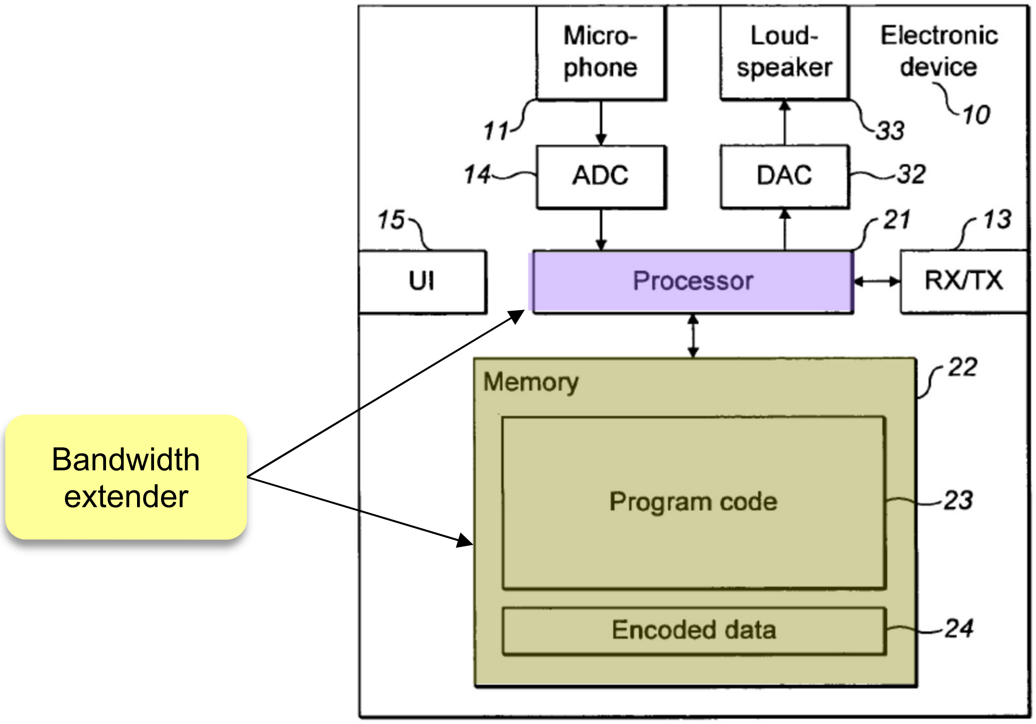


Nilsson, FIG. 5 (Top), and FIG. 7 (Bottom) (Annotated)

II. THE ’060 PATENT

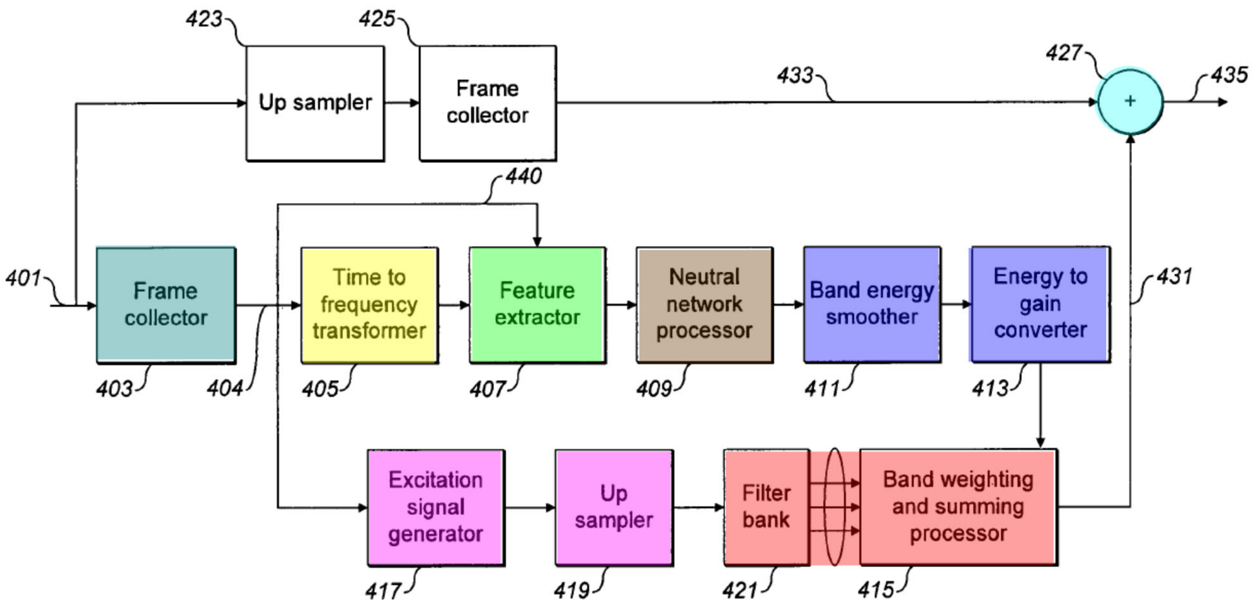
A. ’060 Patent Summary

The ’060 patent describes a method “for extending the bandwidth of an audio signal.” Ex-1001, Abstract.



’060 Patent, FIG. 1 (annotated)

As shown above in annotated FIG. 1, ’060 patent discloses an electronic device 10 including a processor 21 and program code 23 stored in memory 22. The program code 23 “may comprise an audio decoding code or speech decoding code.” Ex-1001, 7:28-29. “The processor 21 may execute the decoding program code stored in the memory 22,” (*Id.*, 4:62-63), and functions as an audio “bandwidth extender.” *Id.,* 8:11-14.



’060 Patent, FIG. 4 (annotated)

As shown above in annotated FIG. 4, the bandwidth extender includes a frame collector 403 “whereby the input audio signal (otherwise known as the audio sample stream) is partitioned and collated into a continual series of audio frames.” *Id.*, 9:36-39.

The bandwidth extender includes a time to frequency transformer 405 “whereby the time based audio signal frame 404 of 96 samples can be transformed to the frequency domain.” *Id.*, 10:6-9.

The bandwidth extender includes a feature extractor 407 whereby “frequency domain and time domain features [are] extracted for each frame of the input audio signal and frequency domain signal.” *Id.*, 10:51-54.

The bandwidth extender includes a neural network processor 409 that can "generate the spectral shape parameters for the artificial high band signal 431." *Id.*, 14:65-67.

The bandwidth extender includes a band energy smoother 411 that can "filter the energy level for each sub band over current and past values," (*Id.*, 17:34-36) and an energy to gain converter 413, "whereby an energy level associated with a particular sub band of the filter bank can [] be converted to a suitable gain factor." *Id.*, 19:1-5.

The bandwidth extender includes an excitation signal generator 417 that generates excitation signal using "the time domain frames" and an up-sampler 419 that up-sample the output excitation signal to generate high band excitation signal. *Id.*, 18:51-56.

The bandwidth extender includes a filter bank 421 and a band weighting and summing processor 415 whereby "each sub band from the filter bank 421 can [ ] be individually weighted by a corresponding sub band gain factor," and the weighted sub band signals are summed together to form the artificially generated high band signal 431. *Id.*, 18:56-67.

"The artificially generated high band signal 431 [] then be passed to an input of a summer 427 in which the signal 431 is combined with an up sampled input audio signal 433 to produce the bandwidth extended signal 435." *Id.*, 26:44-48.

B. ’060 Patent Prosecution History

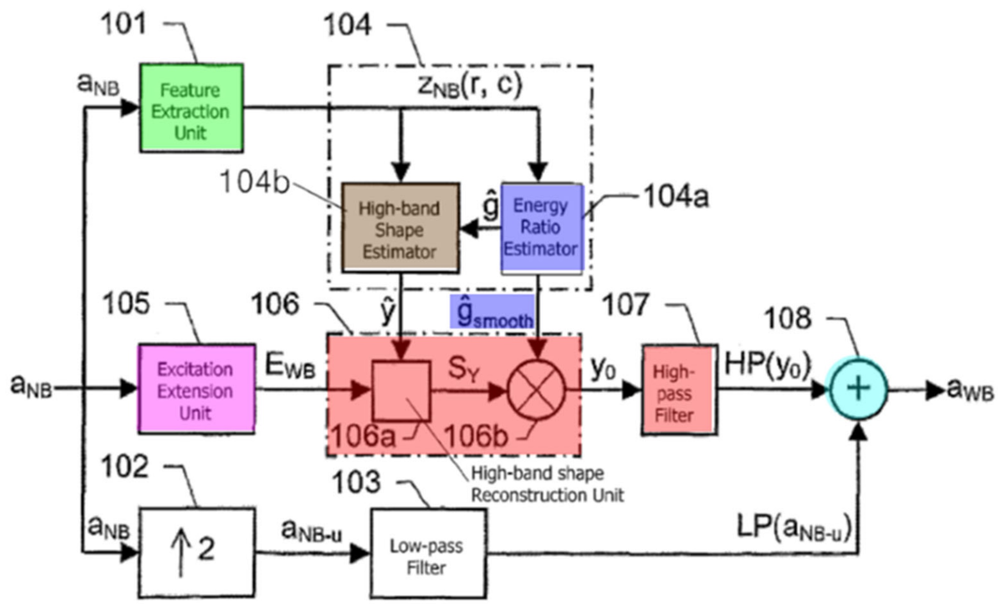
The prosecution history of ’060 patent was surprisingly short. In the first Office Action (dated June 25, 2015), the Examiner allowed dependent claims 2-5, 7, 14-17, and 19. Ex-1002, pages 230-231. In response, Applicant amended independent claims to incorporate the subject matter of allowable claim 7. Ex-1002, pages 241-249.

However, in allowing the claims, the Examiner did not consider any of the references set forth in this Petition. Claims 1-18 of ’060 patent are unpatentable over the references set forth in this Petition.

III. OVERVIEW OF PRIOR ART

A. *Nilsson*

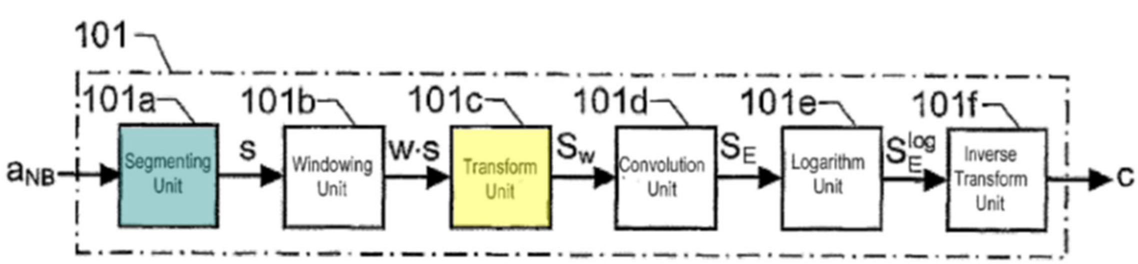
Like ’060 patent, *Nilsson* is directed to a bandwidth extension. As shown below in annotated FIG. 5, *Nilsson* discloses a signal decoder that “receives a narrow-band acoustic signal a_{NB}” and generates “the wide-band acoustic signal a_{WB} [that] has a wider spectrum A_{WB} than the spectrum A_{NB} of the narrow-band acoustic signal a_{NB}.” Ex-1005, 5:54-61; 4:24-36.



Nilsson, FIG. 5 (annotated)

Nilsson's decoder includes a feature extraction unit 101 that "receives the narrow-band acoustic signal a_{NB} and produces in response thereto at least one essential feature Z(r, c) that describes particular properties of the received narrow-band acoustic signal a_{NB}." *Id.,* 6:25-28.

As shown below in annotated FIG. 7, *Nilsson* uses a transform unit 101c included in feature extraction unit 101 and "computes a corresponding spectrum SW by means of a fast Fourier transform." *Id.,* 6:56-58.



Nilsson, FIG. 7 (annotated)

The extracted essential feature Z(r, c) includes time domain component feature such as "[t]he degree of voicing r, which represents one such essential feature zNB(r, c)," (*Id.,* 6:25-32), and frequency domain component feature such as "the spectral envelope c." *Id.,* 6:15-21.

Nilsson's decoder includes a high-band shape estimator 104b "to create a combination of the high-band shape and energy-ratio." *Id.,* 10:31-38.

Nilsson's decoder includes an energy-ratio estimator 104a that "receives the first component c0 in the vector of linear frequency cepstral coefficients c and produces, on basis thereof, plus on basis of the narrow-band shape x and the degree of voicing r an estimated energy-ratio $\hat{g}$ between the high-band and the narrow-band." *Id.,* 7:15-21.

Nilsson's decoder includes an excitation extension unit 105 that "receives the narrow-band acoustic signal a_{NB} and, on basis thereof, produces an extended excitation signal EWB." *Id.,* 10:51-53.

Nilsson's decoder includes a high band shape estimator 106a in the wide-band filter 106 that performs frequency shaping of the excitation signal and a high-pass filter 107 that produces "a high-pass filtered signal HP(y0)." *Id.,* 11:21-59.

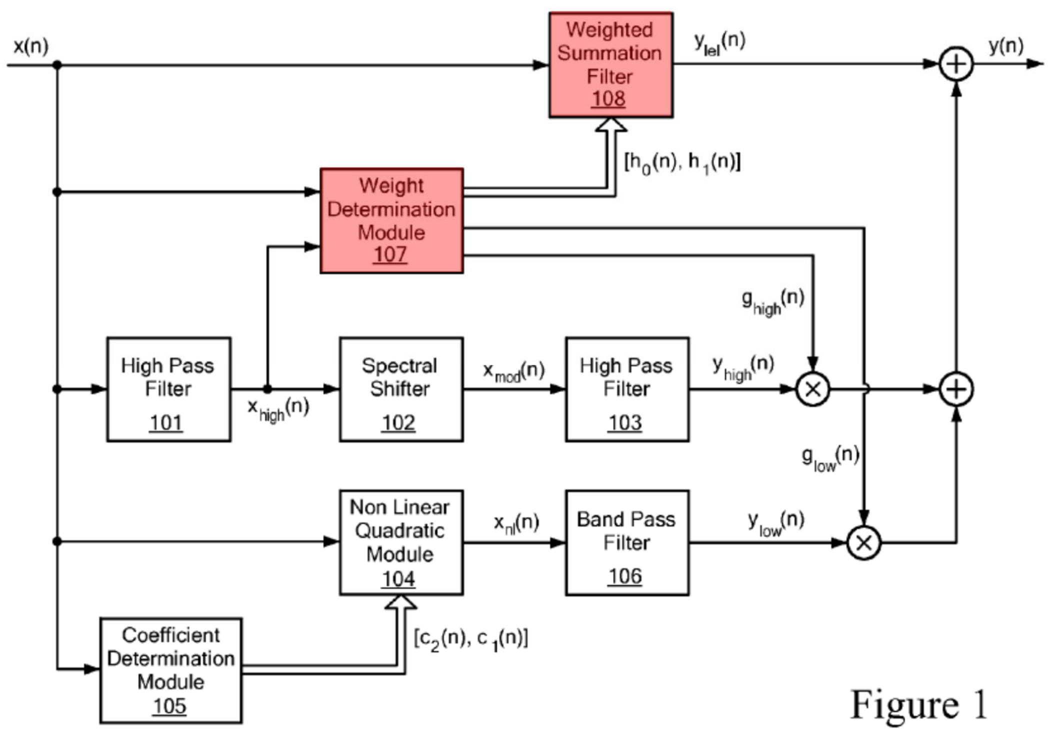
Nilsson's decoder includes an "adder 108 [that] receives the low-pass filtered signal LP(a_{NB}-u), receives the high-pass filtered signal HP(y0) and adds the

received signals together and thus forms the wide-band acoustic signal a_{WB}.” *Id.,* 12:17-21.

B. *Iser*

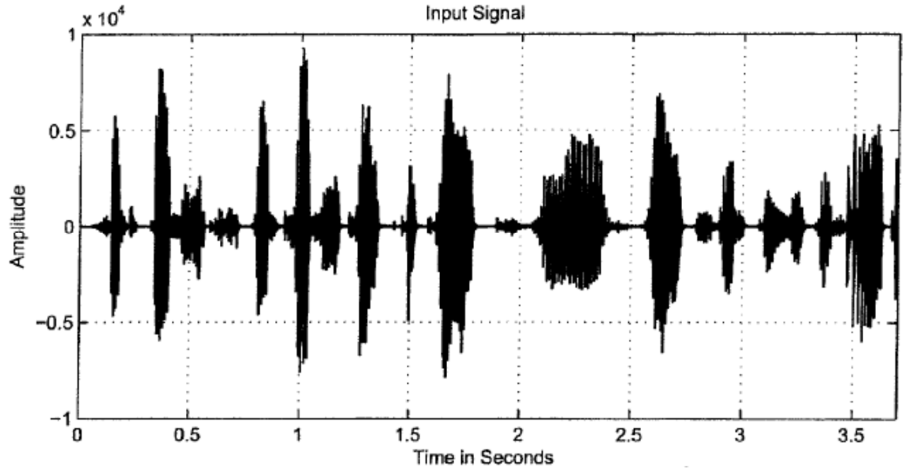
Like ’060 patent, *Iser* discloses “a bandwidth extension system for providing an acoustic signal with extended bandwidth.” Ex-1006, 2:19-24.

As shown below in annotated FIG. 1, *Iser*’s bandwidth extension system includes a weight determination module 107 that determines “weighting factors for the other signal component,” (*Id.*, 9:39-41), and a weighted summation filter 108 which performs “weighted sum of the received acoustic signal x(n) at time n and at time n−1.” *Id.,* 9:38-39.



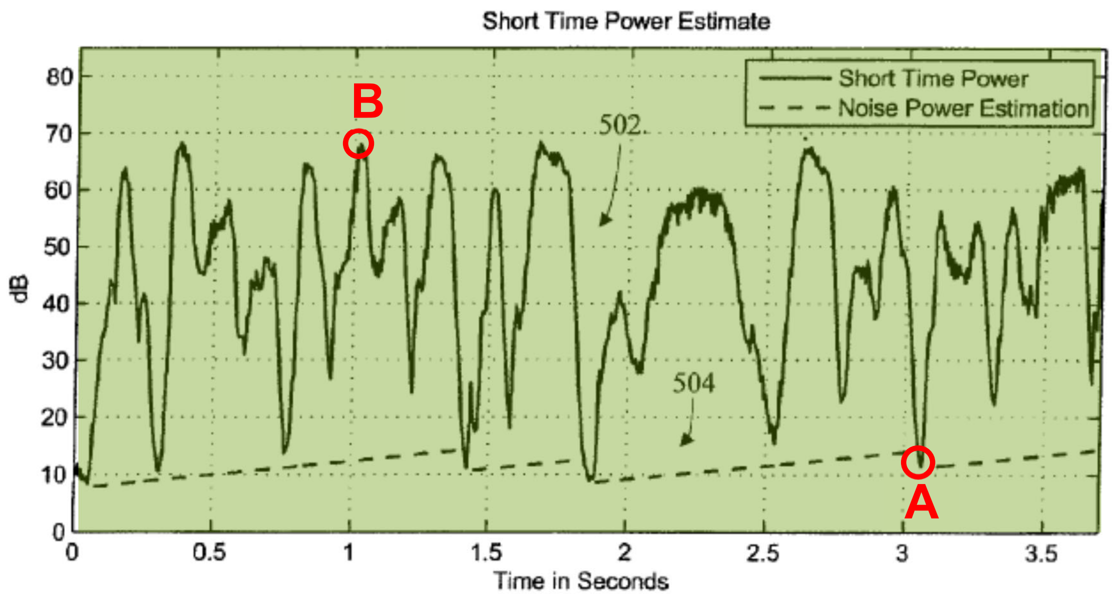
Iser, FIG. 1 (annotated)

In applying the weighting factors, *Iser* compares the power of the input audio signal and the estimated noise power. "FIG. 4 is a representation of an input speech signal, such as the received acoustic signal x(n)." *Id.,* 7:8-10.



Iser, FIG. 4

As shown below in annotated FIG. 5, *Iser* estimates "a short time power and a noise power that correspond to the received acoustic signal x(n) of FIG. 4. Line 502 in FIG. 5 represents the estimated short time power x(n) of the received acoustic signal x(n). Line 504 in FIG. 5 represents the noise power estimation b(n). The short time power estimation may be used to determine different factors for weighting the signal components." *Id.,* 7:8-18.



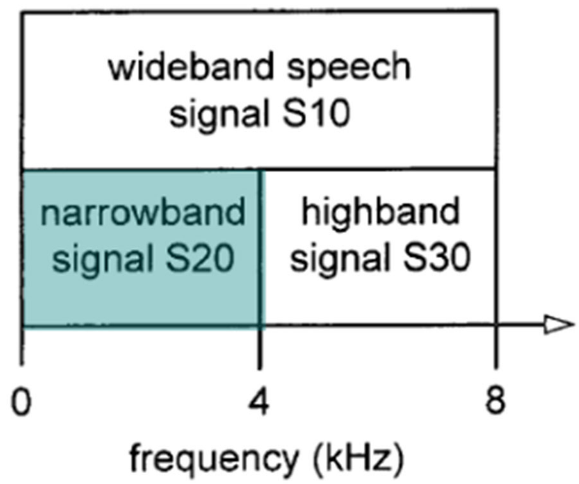
Iser, FIG. 5 (annotated)

Based on the above estimation of signal power and noise power, *Iser* determines "signal-to-noise ratio of the received acoustic signal" and applies a weighting factor as "a function of an estimated signal-to-noise ratio of the received acoustic signal x(n)." *Id.,* 7:20-32. That is, "[i]f the received acoustic signal x(n) contains speech passages with a **low signal-to-noise ratio**, then this weighting factor may be used to **damp** the upper bandwidth extension signal yhigh(n)." *Id.,* 7:20-32. In the above annotated FIG. 5, point A indicates a low signal-to-noise ratio (i.e., signal power approaches the noise power estimation) and point B indicates a high signal-to-noise ratio.

C. *Vos*

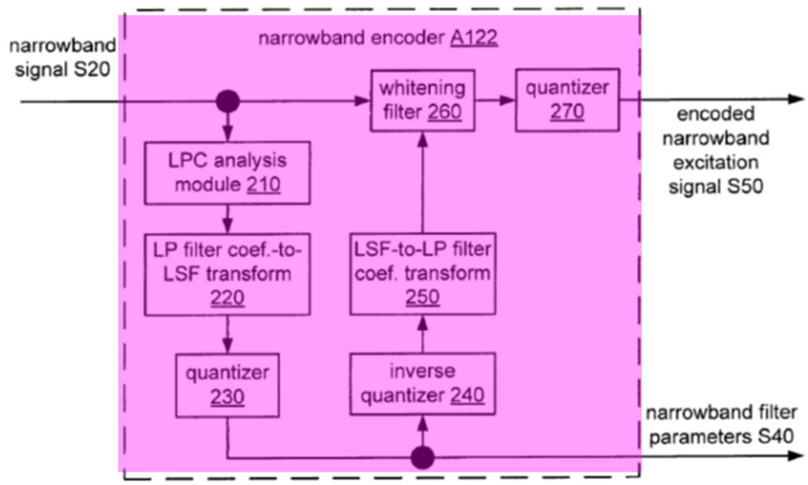
Like '060 patent, *Vos* is directed to bandwidth extension. Ex-1007, [0011].

As shown below in annotated FIG. 4, *Vos* starts with generating a narrowband signal S20 using a wideband signal.



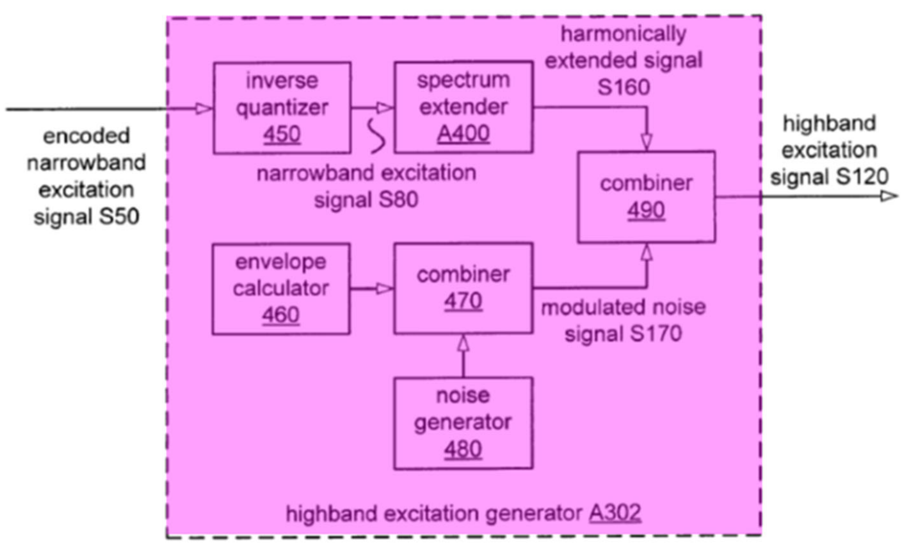
Vos, FIG. 4 (annotated)

Using the generated narrowband signal, *Vos* generates an encoded narrowband excitation signal. As shown below in annotated FIG. 6, a "narrowband encoder A122 generates a residual signal by passing narrowband signal S20 through a whitening filter 260 (also called an analysis or prediction error filter) that is configured according to the set of filter coefficients." *Id.*, [0118].



Vos, FIG. 6 (annotated)

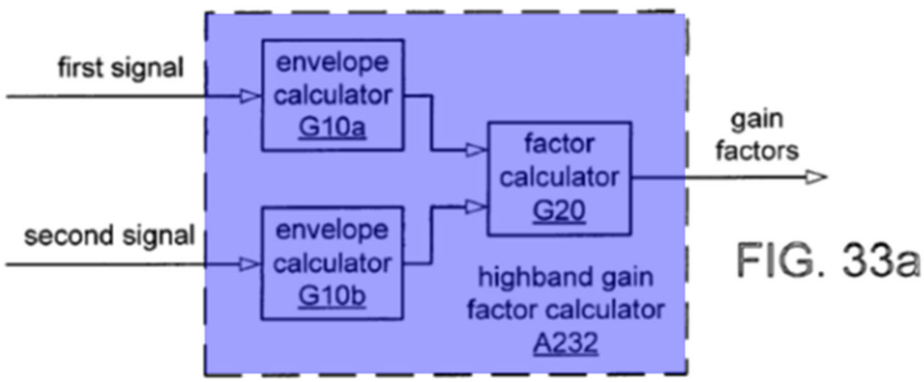
Further, using the encoded narrowband excitation signal, *Vos* generates highband excitation signal. As shown below in annotated FIG. 11, a "[h]ighband excitation generator A300 is configured to generate highband excitation signal S120 by extending the spectrum of narrowband excitation signal S80 into the highband frequency range." *Id.*, [0139].



Vos, FIG. 11 (annotated)

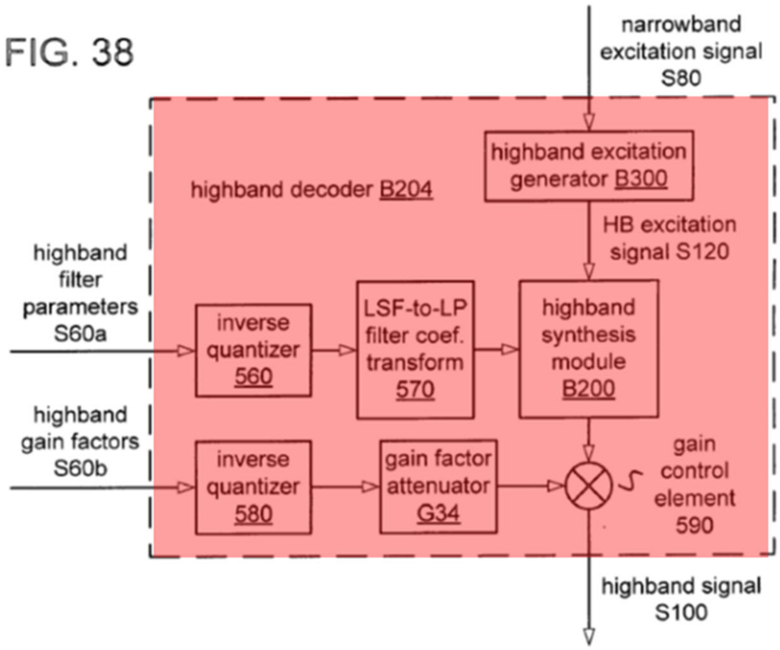
Vos discloses extracting a feature vector including frequency domain feature components, such as spectral tilt, centroid, energy envelope. *Id.*, [0130], [0255], [0218]. *Vos* also discloses extracting a feature vector including time domain feature components, such as voice activity, time-domain envelope. *Id.*, [0131], [0155].

Vos discloses determining spectral shape parameter (e.g., gain factor) using the extracted feature vector. As shown below in annotated FIG. 33a, *Vos* discloses "a highband gain factor calculator A230 that is configured to calculate a series of gain factors." *Id.,* [0216]-[0217].



Vos, FIG. 33a (annotated)

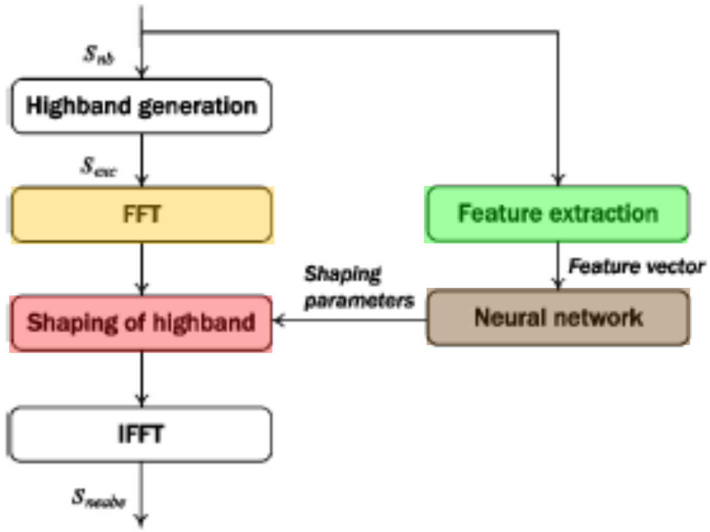
As shown below in annotated FIG. 38, *Vos* discloses a "Highband synthesis filter B200 [that] is configured to produce a synthesized highband signal according to highband excitation signal S120 and the set of filter coefficients." *Id.,* [0181].



Vos, FIG. 38 (annotated)

D. *Kontio*

Like ’060 patent, *Kontio* is directed to bandwidth extension. *Kontio*’s method is neural network-based artificial bandwidth expansion of speech. As shown below in annotated FIG. 1, in *Kontio*’s method “the narrowband input signal, denoted by Snb, is treated in time-domain frames. New frequency components are produced into the highband using spectral folding, which can be achieved, for example, by zero-insertion in the time-domain. The frame is then transformed into the frequency domain with FFT. A set of features is calculated from the narrowband signal and given as an input to a **neural network** which transforms them into shaping parameters.” Ex-1008, page 874, 2:18-25.



Kontio, FIG. 1 (annotated)

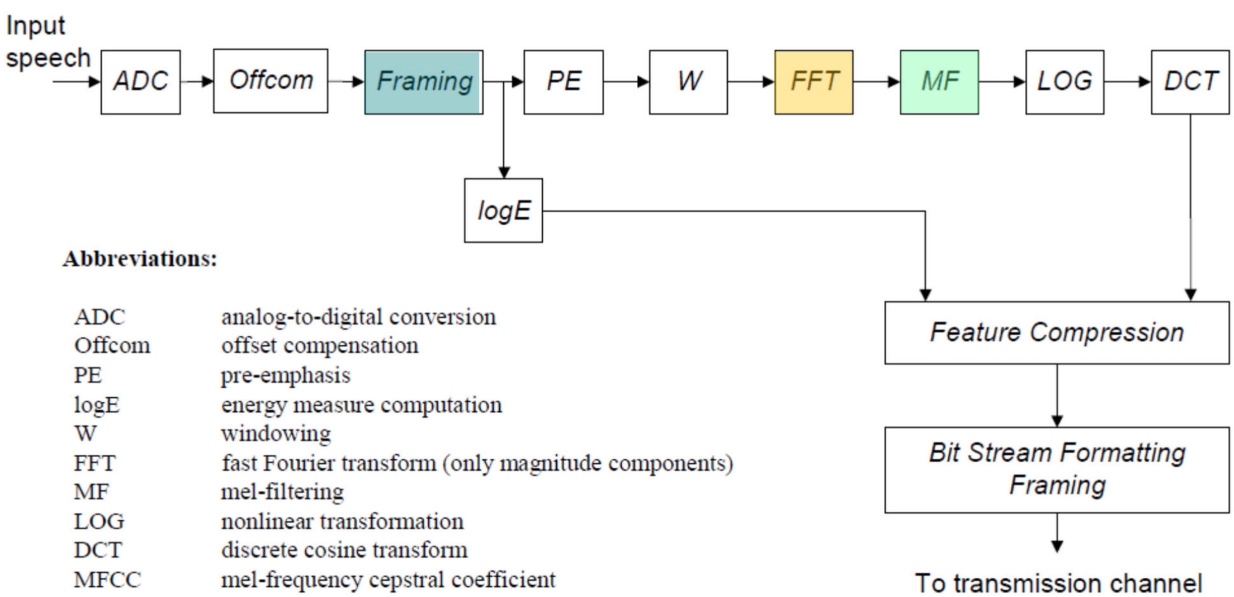
Kontio also discloses extracting feature vector such as **gradient index**, differential energy ratio, frame energy ratio, and subband power levels. *Id.*, page 875, 1:35-2:19. For the subband power level estimation, *Kontio* selects "the average spectral magnitudes of four spectral subbands (0.3–1.0 kHz, 1.0–1.7 kHz, 1.7–2.5 kHz, and 2.5–3.4 kHz)." *Id.*, page 875, 2:12-15. *Kontio* also discloses that the neural network is **trained** to provide shaping parameters for shaping highband signal. *Id.*, Sections II-B and II-C.

E. *ETSI201.108*

ETSI201.108 was released by European Telecommunications Standards Institute (ETSI) in 2003 and publicly available on the ETSI server from September

23, 2003. Ex-1010[2]. No particular indica is *per se* sufficient for satisfying the "reasonable likelihood" publicly available standard, but for a given form of technology, screenshots from internet archive sources, and evidence of printed version release dates are considered among the examples of sufficient indica evidencing public accessibility. *Hulu, LLC v. Sound View Innovations, LLC,* Case IPR2018-013039, Paper 29 (P.T.A.B. Dec. 20, 2019) (precedential). For instance, the ETSI website demonstrates that *ETSI201.108* was published on September 23, 2003—a screen shot of the ETSI website showing the publication date is submitted as Ex-1010.

ETSI201.108 discloses "the algorithm for front-end feature extraction to create Mel-Cepstrum parameters." *ETSI201.108*, page 5, line 4.



[2] *See also* the Internet Archive stored document at https://web.archive.org/web/20031215120853/https://www.etsi.org/.

ETSI201.108, FIG. 4.1 (annotated)

As shown above in annotated FIG. 4.1, an input speech signal goes through an analog-to-digital conversion step and an offset compensation step, and then "is divided into **overlapping frames** of N samples" in a framing step. Ex-1009, page 9, line 24.

After the framing, the signal is divided into two paths. The first path is for logarithmic frame energy measure (logE). In the second path, the signal goes through a pre-emphasis step by a pre-emphasis filter and a windowing step by Hamming window, and then converted to frequency domain signal by fast Fourier transform (FFT).

The useful band of the converted frequency band "is divided into 23 channels **equidistant** in **mel frequency** domain. Each channel has **triangular-**shaped frequency window. Consecutive channels are half-**overlapping**." *Id.*, page 10, line 25-27. "The output of mel filtering is subjected to a logarithm function (natural logarithm)" at step LOG and then 13 Mel-frequency cepstral coefficients (MFCC) are calculated from the output of the Non-linear Transformation block" at step DCT. *Id.*, page 10, line 13-20.

The final feature vector consists of 14 coefficients: the log-energy coefficient and the 13 Mel-frequency cepstral coefficients. The feature vectors are compressed and formatted as bitstream for further transmission. Thus, Mel

frequency cepstral coefficients (MFCC) are "cepstral coefficients calculated from the mel-frequency warped **Fourier transform representation** of the log magnitude spectrum." *Id.*, page 6, lines 23-26.

IV. LEVEL OF ORDINARY SKILL IN THE ART

The level of ordinary skill in the art may be reflected by the prior art of record. *See Okajima v. Bourdeau*, 261 F.3d 1350, 1355 (Fed. Cir. 2001). A person of ordinary skill in the art ("POSITA") for '060 patent would at least have a bachelor's degree in electrical engineering, computer engineering, or a related engineering discipline and two or more years of industry experience in the field of signal processing, or equivalent experience, education, or both. Ex-1004. The person would also have knowledge or familiarity with signal processing. *Id.*

V. CLAIM CONSTRUCTION

Only claim terms "in controversy" need be construed in IPR "and only to the extent necessary to resolve the controversy." *Nidec Motor Corp. v. Zhongshan Broad Ocean Motor Co.*, 868 F.3d 1013, 1017 (Fed. Cir. 2017) (citation omitted), *cert. denied*, 138 S. Ct. 1695 (2018). No claim terms need to be construed by the Board at this time.

In the District Court, Petitioner and Patent Owner have offered the following terms for construction, *see* Litigation, Dkt. 61, p. 5, after the number of terms have been narrowed:

Claim	Term	Petitioner's Proposed Construction	Patent Owner's Proposed Construction
1 and 10	"extracting a feature vector"	Indefinite	Plain and Ordinary Meaning
1 and 10	"the level value is attenuated"	Indefinite	Plain and Ordinary Meaning
1 and 10	"spectral shape parameter"	A sub band energy level value or a sub band gain factor based on the sub band energy level value	Plain and Ordinary Meaning

For the purpose of this *Inter Partes* Review, Petitioner adopts Patent Owner's constructions for the above terms. Nevertheless, the evidence presented in this Petition is sufficient to invalidate the claims without addressing the indefinite issue.

VI. STATEMENT OF PRECISE RELIEF REQUESTED FOR EACH CLAIM CHALLENGED

1. Claims for Which Review Is Requested

Petitioners respectfully request review under 35 U.S.C. § 311 of claims 1-18 of '060 patent and cancellation of those claims as unpatentable.

2. Statutory Grounds

Each asserted reference identified in the table below issued, published, and/or was filed before May 25, 2010, the earliest purported priority date of '060 patent. Thus, each asserted reference is prior art under at least one of pre-AIA 35

U.S.C. §§ 102(a), (b), and/or (e).

Prior Art References	Exhibit
Nilsson, U.S. Patent No. 7,359,854, filed on Apr. 10, 2002, issued on Apr. 15, 2008.	Ex-1005
Iser, U.S. Patent No. 8,160,889, filed on Jan. 17, 2008, issued on Apr. 17, 2012.	Ex-1006
Vos, U.S. Patent App. Pub. No. 2006/0282262, filed on Apr. 21, 2006, published on Dec. 14, 2006.	Ex-1007
Kontio, J. *Kontio*, L. Laaksonen and P. Alku, "Neural Network-Based Artificial Bandwidth Expansion of Speech," in *IEEE Transactions on Audio, Speech, and Language Processing,* vol. 15, no. 3, pp. 873-881, March 2007.	Ex-1008
ETSI201.108, ETSI ES 201 108 V1.1.3 (2003-09), ETSI Standard, Speech Processing, Transmission and Quality Aspects (STQ); Distributed speech recognition; Front-end feature extraction algorithm; Compression algorithms, published in 2003.	Ex-1009

Claims 1-18 of ’060 patent are unpatentable under the following grounds based on pre-AIA 35 U.S.C. § 103:

Grounds	Grounds of Unpatentability
1	*Nilsson* and *Iser* render obvious claims 1, 9, 10, and 18.
2	*Nilsson*, *Iser*, and *ETSI201.108* render obvious claims 4, 5, 13, and 14
3	*Nilsson*, *Iser*, and *Vos* render obvious claims 2, 3, 7, 11, 12, and 16
4	*Nilsson*, *Iser*, and *Kontio* render obvious claims 6, 8, 15, and 17.

VII. GROUND 1: *NILSSON* AND *ISER* RENDER OBVIOUS CLAIMS 1, 9, 10, AND 18

A. POSITA Would Have Combined *Nilsson* and *Iser*

Both *Nilsson* and *Iser* are directed to a method for bandwidth extension to generate bandwidth extended audio signal using a narrow-band audio signal in a frequency range of 0-4 kHz. Ex-1003, ¶¶76-137.

Each reference includes similar modules with similar functions, for instance, both *Nilsson* and *Iser* disclose adjusting the high-band (e.g., 4-8 kHz) signal components. While *Nilsson* discloses a broad concept of adjusting the high-band energy, *Iser* discloses a specific way of adjusting, that is, by estimating signal power and noise power in the audio signal and determining a signal-to-noise ratio based on the power estimation and applying a weight factor corresponding to the signal-to-noise ratio. POSITA would have motivated to combine *Nilsson* and *Iser*

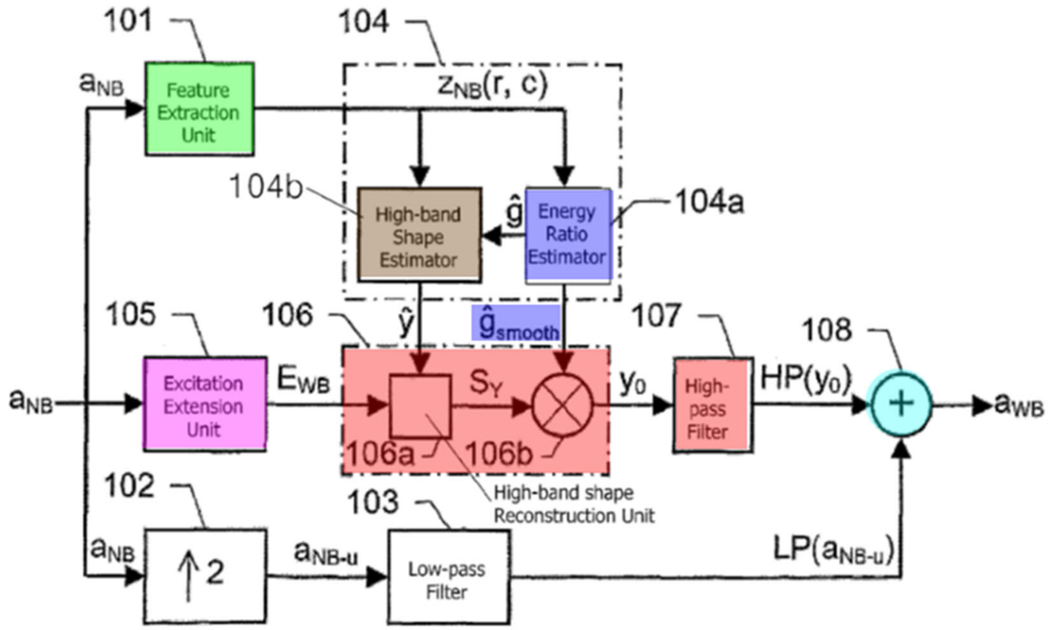
by adopting the specific method of adjusting the high-band energy, as taught by *Iser*, in *Nilsson*'s system, as discussed in detail below.

B. Independent Claim 1

1. [1preamble] "A method comprising:"

Nilsson discloses a method for "extending the spectrum of a received narrow-band acoustic signal (a_{NB}). A wide-band acoustic signal (AWB) is produced by extracting at least one essential attribute (z_{NB}) from the narrow-band acoustic signal (a_{NB}). Parameters, e.g., representing signal energies, with respect to wide-band frequency components outside the spectrum (ANB) of the narrow-band acoustic signal (a_{NB}), are estimated based on the at least one essential attribute (z_{NB})." Ex-1005, Abstract. Ex-1003, ¶¶78-79.

As shown below in annotated FIG. 5, *Nilsson*'s method is performed by a "signal decoder" that "receives a narrow-band acoustic signal a_{NB}" to generate bandwidth extended signal a_{WB}. *Id.*, 5:54-61. Ex-1003, ¶79.



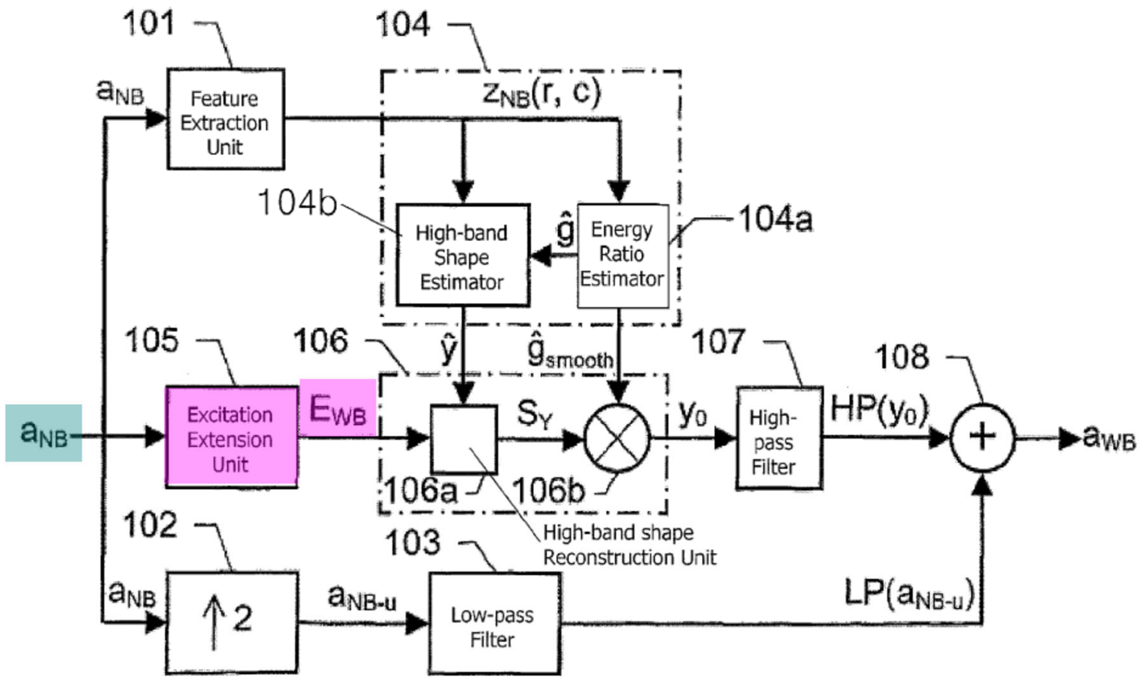
Nilsson, FIG. 5 (annotated)

2. [1a] "generating an excitation signal from an audio signal, wherein in the audio signal comprises a plurality of frequency components;"

Nilsson discloses generating an excitation signal from an audio signal, as claimed, (*Nilsson*'s "excitation extension unit 105" generates "excitation signal EWB" from the "narrow-band acoustic signal a_{NB}"), wherein the audio signal comprises a plurality of frequency components. Ex-1003, ¶¶80-85.

i. *Nilsson* discloses generating an excitation signal from an audio signal.

As shown below in annotated FIG. 5, *Nilsson*'s decoder includes an "excitation extension unit 105 [that] receives the narrow-band acoustic signal a_{NB} and, on basis thereof, produces an extended excitation signal EWB." Ex-1005, 10:51-53.



Nilsson, FIG. 5 (annotated)

Specifically, "the extended excitation signal EWB is generated by means of spectral folding of a corresponding excitation signal ENB for the narrow-band acoustic signal a_{NB} around a particular frequency." *Id.*, 10:57-11:2.

ii. *Nilsson*'s audio signal comprises a plurality of frequency components.

The '060 patent fails to define "a plurality of frequency components" of the audio signal. The '060 patent merely discloses using the low-band audio signal with "effective frequency range from 250 to 3500 Hz" to generate the excitation signal. Ex-1001, 11:9-10. In fact, it is well-known in the art that any audio signal naturally comprises a plurality of frequency components. Ex-1003, ¶83.

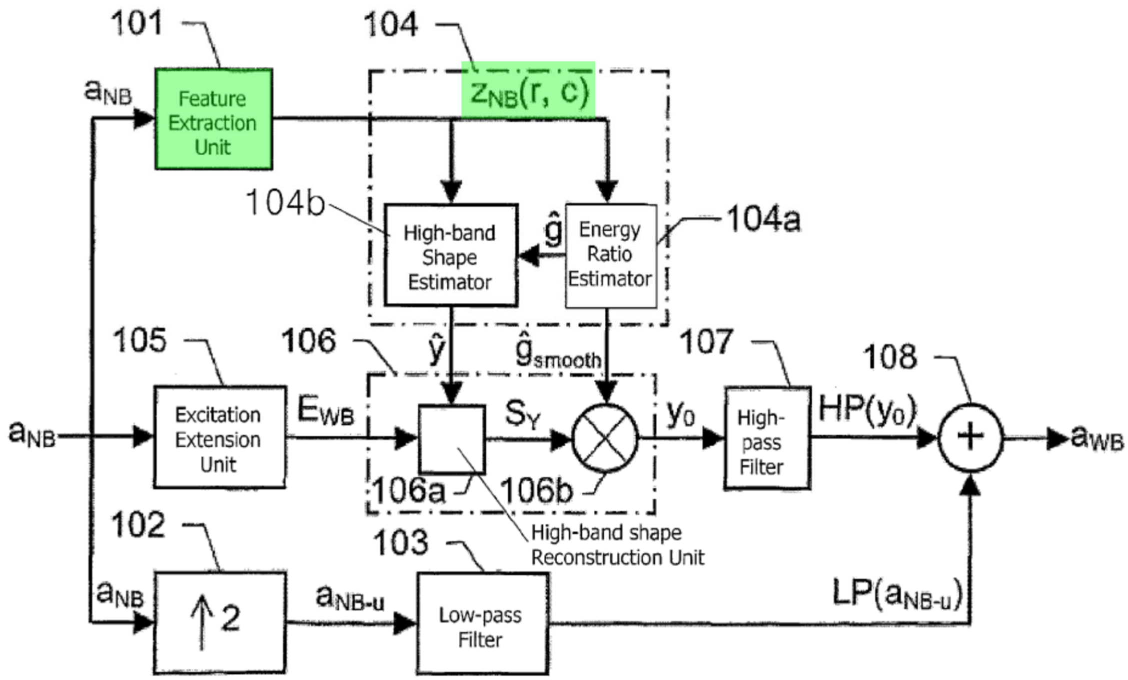
Similarly, *Nilsson*'s narrowband signal includes a plurality of frequency components. In generating the excitation signal, *Nilsson* uses a narrowband signal a_{NB} which is a frequency band that covers "channel bandwidth of today's public switched telephony networks (PSTNs)", for example, "between 0.3 kHz and 3.4 kHz" (Ex-1005, 1:16-21) or "between 0 kHz and 4 kHz." Ex-1005, 1:55-58. Ex-1003, ¶84.

Therefore, *Nilsson* discloses generating an excitation signal from an audio signal, wherein in the audio signal comprises a plurality of frequency components. Ex-1003, ¶85.

3. [1b] "extracting a feature vector from the audio signal, wherein the feature vector comprises at least one frequency domain component feature and at least one time domain component feature;"

Nilsson discloses extracting a feature vector from the audio signal, as claimed, (*Nilsson*'s "feature extraction unit 101" extracts the "essential feature z_{NB}(r, c)" from "narrow-band acoustic signal a_{NB}"), wherein the feature vector comprises at least one frequency domain component feature, as claimed, (e.g., "spectral envelope c" of *Nilsson*) and at least one time domain component feature, as claimed, (e.g., "degree of voicing r" of *Nilsson*). Ex-1003, ¶¶86-94.

As shown below in annotated FIG. 5, *Nilsson*'s decoder includes "feature extraction unit 101." Ex-1005, 4:45-46.



Nilsson, FIG. 5 (annotated)

Nilsson discloses that "[t]he narrow-band acoustic signal a_{NB} is fed in parallel to the feature extraction unit 101, the excitation extension unit 105 and the up-sampler 102." *Id.,* 4:55-58. "The feature extraction unit 101 receives the narrow-band acoustic signal a_{NB} and produces in response thereto at least one essential feature $z_{NB}(r, c)$ that describes particular properties of the received narrow-band acoustic signal a_{NB}." *Id.,* 6:25-28.

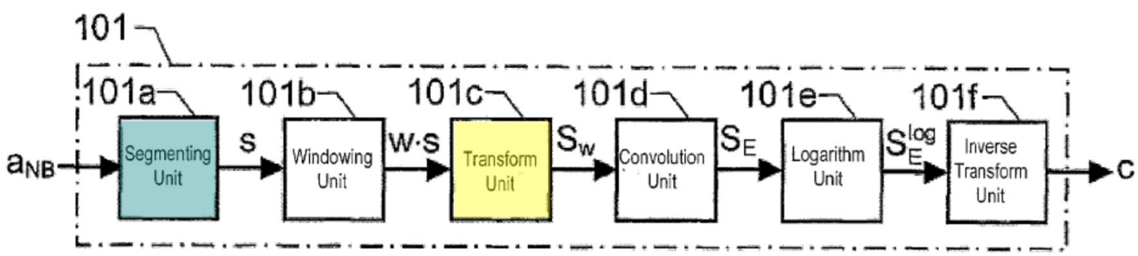
i. Time-domain component feature

Nilsson discloses that the essential feature $z_{NB}(r, c)$ includes "[t]he degree of voicing r which represents one such essential feature $z_{NB}(r, c)$." *Id.,* 6:28-32. The degree of voicing r of *Nilsson* is a time-domain feature because it "is determined by localising a maximum of a normalised autocorrelation function," using sampled

"narrow-band acoustic segment having a duration of Tf (e.g. 20 ms)," without converting the narrowband time-domain signal into frequency-domain signal. *Id.*, 6:30-46. Ex-1003, ¶89.

ii. Frequency-domain component feature

Nilsson's essential feature z_{NB}(r, c) includes a "spectral envelope c." *Id.,* 6:46-51. The spectral envelop c is a frequency-domain component feature because spectral envelop c is determined by converting the time-domain narrowband input signal into a frequency-domain signal by Fourier transform. As shown below in annotated FIG. 7, *Nilsson* discloses "a part of the feature extraction unit 101, which is utilised for determining the spectral envelope c." *Id.*, 6:48-51. Ex-1003, ¶90.



Nilsson, FIG. 7 (annotated)

The part of the feature extraction unit 101 which is utilized for determining the spectral envelope c includes "[a] segmenting unit 101a [that] separates a segment s of the narrow-band acoustic signal a_{NB} that has a duration of Tf=20 ms," and "a transform unit 101c [that] computes a corresponding spectrum SW by means of a **fast Fourier transform**, i.e. Sw=FFT(w·s)," that converts a time-domain signal into a frequency-domain signal. *Id.*, 6:56-58.

Nilsson discloses that the spectral envelope c is "represented by LFCCs [linear frequency cepstral coefficients]." *Id.*, 6:47, 5:29-30. Alternatively, the spectral envelope c is represented by "Line Spectral Frequencies (LSF), **Mel Frequency Spectral Coefficients (MFCC)**, and Linear Prediction Coefficients (LPC)." *Id.*, 8:45-50.

Among these alternative representations of the spectral envelope c, at least Mel frequency cepstral coefficients (MFCC) are frequency domain component feature because it is well-known in the art that Mel frequency cepstral coefficients are "cepstral coefficients calculated from the mel-frequency warped **Fourier transform representation** of the log magnitude spectrum." See *ETSI201.108*, page 6, lines 23-26. *See* also above Section III.E. Ex-1003, ¶93.

Therefore, *Nilsson* discloses extracting a feature vector from the audio signal, wherein the feature vector comprises at least one frequency domain component feature and at least one time domain component feature. Ex-1003, ¶94.

4. [1c] "determining at least one spectral shape parameter from the feature vector, wherein the at least one spectral shape parameter corresponds to a sub band signal comprising frequency components which belong to a further plurality of frequency components; and"

Nilsson discloses determining at least one spectral shape parameter from the feature vector, as claimed, (*Nilsson* discloses determining energy-ratio ĝ from the essential feature z_{NB}(r, c)), wherein the at least one spectral shape parameter

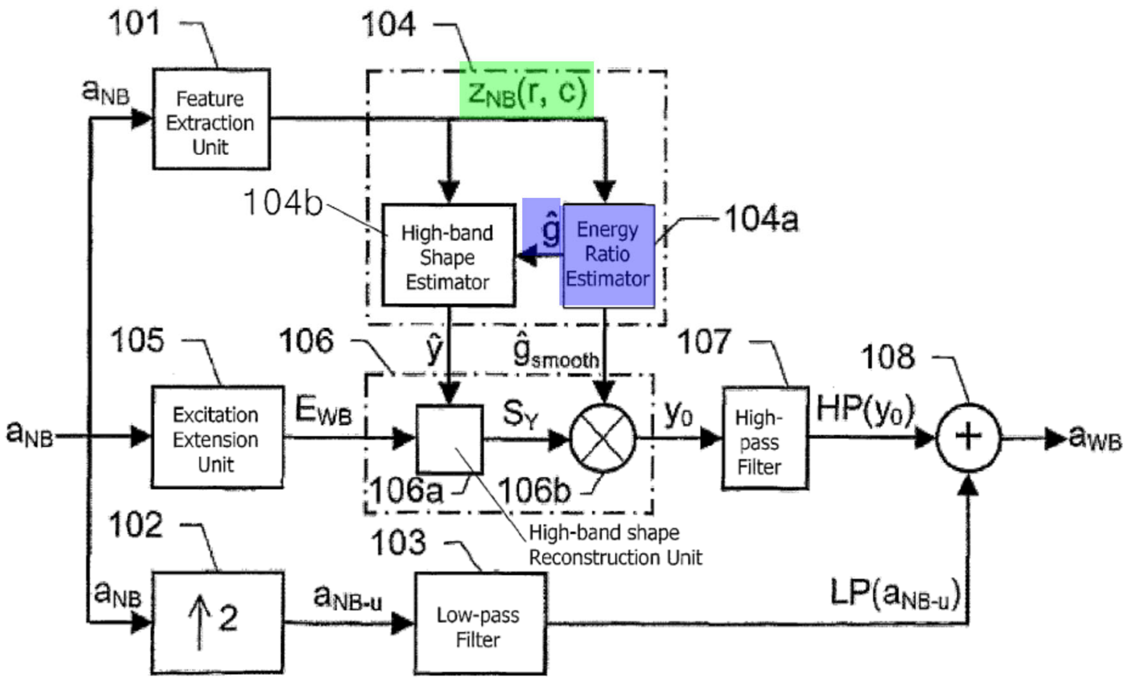
corresponds to a sub band signal comprising frequency components which belong to a further plurality of frequency components. Ex-1003, ¶¶95-104.

i. *Nilsson* discloses determining at least one spectral shape parameter from the feature vector.

The ’060 patent discloses that “[t]he spectral shape parameter may be a sub band **gain factor** based on the sub band energy level value.” Ex-1001, 5:5-8 (emphasis added).

Nilsson similarly discloses determining energy-ratio $\hat{g}$ from the essential feature $z_{NB}(r, c)$. It is well-known in the art that a gain factor indicates a relative change of signal parameters (e.g., power, current, voltage, energy, etc.) and is usually expressed as a linear ratio of the parameters or a logarithm difference of the parameters as in *Nilsson*. Ex-1003, ¶97. Therefore, the energy-ratio $\hat{g}$ of *Nilsson* is or corresponds to the gain factor. Ex-1003, ¶97.

As shown below in annotated FIG. 5, *Nilsson*’s decoder includes an “energy-ratio estimator 104a, which is included in the wide-band envelope estimator 104, receives the first component c0 in the vector of linear frequency cepstral coefficients c and produces, on basis thereof, plus on basis of the narrow-band shape x and the degree of voicing r an estimated energy-ratio $\hat{g}$ between the high-band and the narrow-band.” Ex-1005, 7:15-27.



Nilsson, FIG. 5 (annotated)

As discussed above, *Nilsson*'s energy-ratio ĝ is determined by computing linear frequency cepstral coefficients (LFCC), which is merely one type of representation of spectral envelope c. *Nilsson* also discloses that "[o]ther parameters than LFCCs can be used as alternative representations of the narrow-band spectral envelope x. Line Spectral Frequencies (LSF), Mel Frequency Spectral Coefficients (MFCC), and Linear Prediction Coefficients (LPC) constitute such alternatives." *Id.*, 8:45-50. Therefore, *Nilsson* discloses determining energy-ratio ĝ by computing Mel Frequency Spectral Coefficients (MFCC), and thus, all inherent features associated with MFCC. Ex-1003, ¶99.

ii. *Nilsson*'s spectral shape parameter corresponds to a sub band signal comprising frequency

components which belong to a further plurality of frequency components.

The ’060 patent fails to define the term of “belong to a further plurality of frequency components.”

Regarding the “sub band signal”, ’060 patent discloses that “[t]he apparatus may further comprise a signal combiner configured to combine the sub band signal with the audio signal to provide a bandwidth extended audio signal.” Ex-1001, 6:33-36. Therefore, the claimed “sub band signal” is the high-band signal. Ex-1003, ¶101.

As discussed above, *Nilsson* discloses determining the spectral shape parameter for the high-band signal. Ex-1003, ¶102.

Nilsson also discloses that the narrow-band shape is “modelled by the cepstral coefficients.” Ex-1005, 6:7-8. As discussed above, the cepstral coefficient can be represented by linear frequency cepstral coefficients (LFCC), or alternatively, by Mel Frequency Spectral Coefficients (MFCC). *See Id.*, 8:45-50. It is well-known in the art that in computing MFCC, the frequency band is divided into a plurality of channels and each channel including a plurality of frequency component. *See* Ex-1009, *see* also Section III.E. Thus, *Nilsson*’s cepstral coefficients correspond to a plurality of frequency components, which belong to a further plurality of frequency components. Ex-1003, ¶103. Moreover, the term “sub bands” is well-known in the art. *See* Ex-1008, *see* also Section III.D.

Therefore, *Nilsson* discloses determining at least one spectral shape parameter from the feature vector, wherein the at least one spectral shape parameter corresponds to a sub band signal comprising frequency components which belong to a further plurality of frequency components. Ex-1003, ¶104.

5. [1d] "generating the sub band signal by filtering the excitation signal through a filter bank and weighting the filtered excitation signal with the at least one spectral shape parameter,"

As discussed above in claim element [1c], the claimed "sub band signal" is the high-band signal.

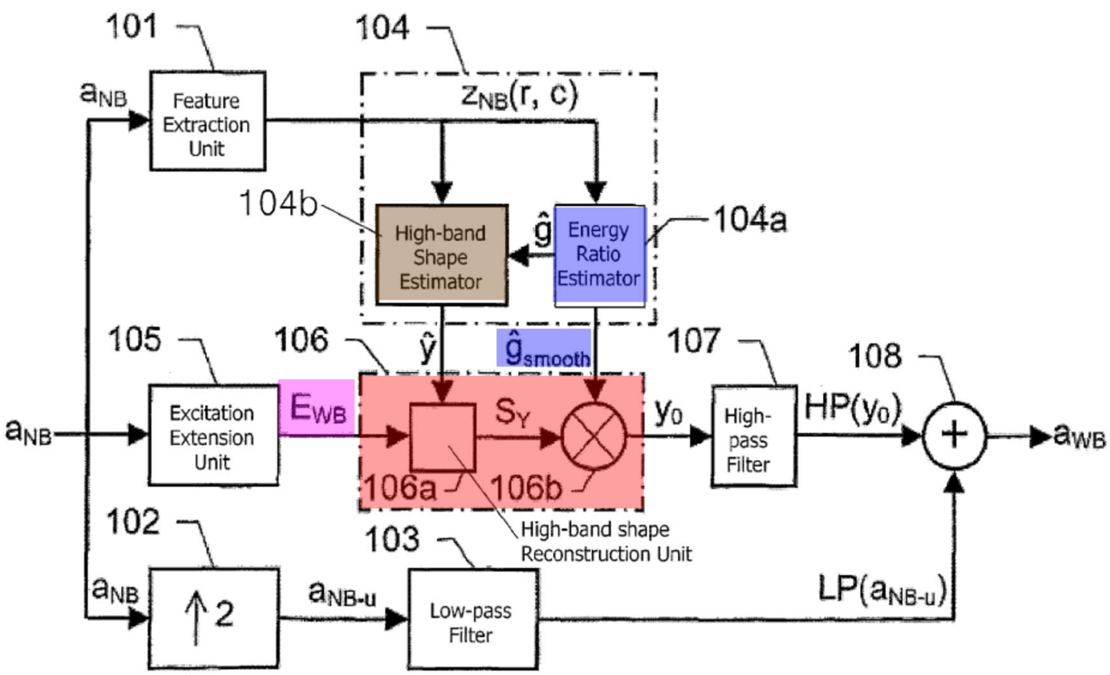
Nilsson discloses generating the high-band signal by filtering the excitation signal through a filter bank and weighting the filtered excitation signal with the at least one spectral shape parameter, as claimed, (*Nilsson* discloses generating high-band signal by filtering the wide-band excitation spectrum EWB through the wide-band filter 106 and multiplying the filtered excitation signal with the smoothed energy ratio ĝ). Ex-1003, ¶¶106-113.

i. *Nilsson* discloses filtering the excitation signal with a filter bank.

The '060 patent does not define the "filter bank."

Nevertheless, *Nilsson* discloses generating the high band signal by filtering the wide-band excitation spectrum EWB through the wide-band filter 106. For instance, as shown below in annotated FIG. 5, *Nilsson* discloses that the "high

band shape estimator 106a in the wide-band filter 106 [] receives the wide-band excitation spectrum EWB from the excitation extension unit 105" to filter the excitation signal. Ex-1005, 11:21-25.



Nilsson, FIG. 5 (annotated)

Also, filtering an excitation signal with a filter bank is well-known in the art. *See* e.g., Ex-1007 (in which *Vos* uses a synthesis filter 330 to filter an excitation signal, which can be implemented as a filter bank). Ex-1003 ¶109. Further, the wide-band filter 106 of *Nilsson* can be implemented as a filter bank. Therefore, the wide-band filter 106 of *Nilsson* corresponds to the claimed filter bank. Ex-1003 ¶109.

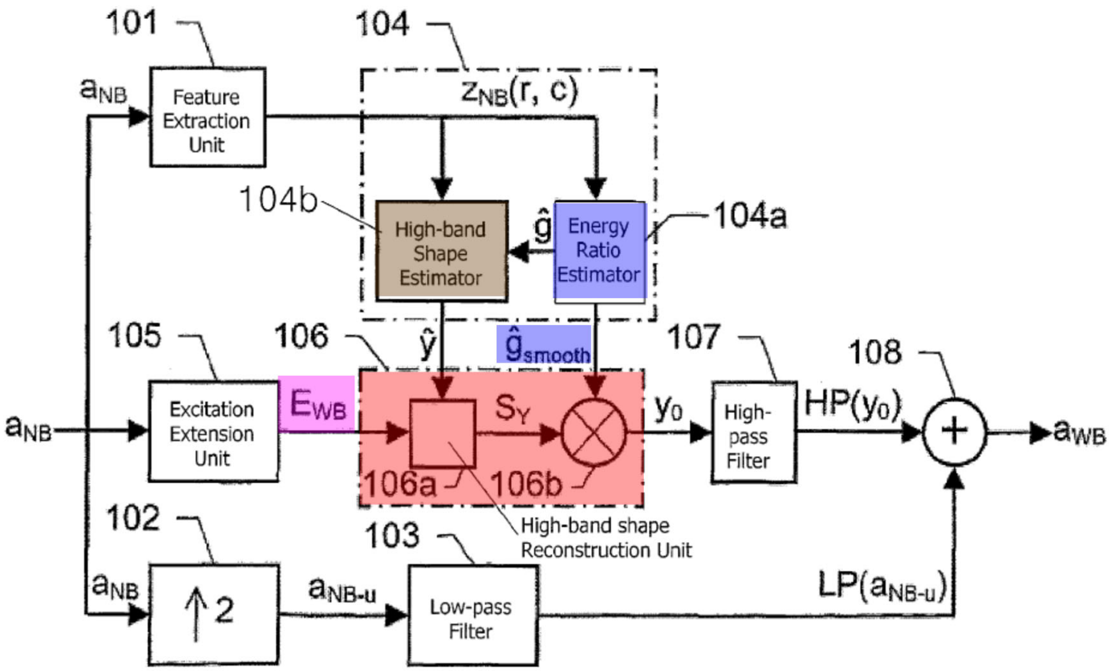
ii. ’060 patent’s disclosure of weighting the filtered excitation signal with the at least one spectral shape parameter.

The ’060 patent only discloses weighting the filtered excitation signal with the gain factor. Specifically, ’060 patent discloses “inputting the time domain frames into an excitation signal generator 417 up-sampling the output of the excitation signal generator 417 in the up-sampler 419 filtering an up-sampled excitation signal through the filter bank 421 and then weighting each sub band signal with a **gain factor** derived from the corresponding mel band energy levels.” Ex-1001, 18:51-58 (emphasis added).

Therefore, “the at least one spectral shape parameter” according to claim element [1d] is the **gain factor**. Ex-1003 ¶111.

iii. *Nilsson* discloses weighting the filtered excitation signal with the at least one spectral shape parameter.

Nilsson discloses weighting the filtered excitation signal with energy-ratio $\hat{g}$ (the claimed spectral shape parameter). As shown below in annotated FIG. 5, *Nilsson* discloses that the “multiplier 106b receives the high-band envelope spectrum S_Y from the high band shape estimator 106a and receives the temporally smoothed energy ratio estimate $\hat{g}_{smooth}$ from the energy ratio estimator 104a. On basis of the received signals S_Y and $\hat{g}_{smooth}$ the multiplier 106b generates a high-band energy y0.” Ex-1005, 11:21-44.



Nilsson, FIG. 5 (annotated)

Because the multiplier 106b multiples the filtered excitation signal E_{WB} with the smoothed energy ratio $\hat{g}_{smooth}$ and the high-band envelope spectrum S_Y which is generated based on energy-ratio $\hat{g}$, *Nilsson* discloses weighting the filtered excitation signal with the energy-ratio $\hat{g}$. "The high-pass filter 107 receives the high-band energy signal y0 from the high-band shape reconstruction unit 106 and produces in response thereto a high-pass filtered signal HP(y0)." *Id.,* 11:56-59. Ex-1003 ¶113.

6. [1e] "wherein the spectral shape parameter is a sub band energy level value and the sub band energy level value is attenuated when the power of the audio signal approaches an estimate of the level of noise in the audio signal."

As discussed above, the "spectral shape parameter" according to Claim [1d] is the "**gain factor**" because '060 patent only discloses weighting the filtered excitation signal with the gain factor. Ex-1003, ¶¶114-126.

Nevertheless, *Nilsson* discloses the sub band energy level value, and *Iser* discloses that the sub band energy level value is attenuated when the power of the audio signal approaches an estimate of the level of noise in the audio signal.

i. *Nilsson* discloses sub band energy level value.

As discussed above, the "sub band signal" is the high-band signal. Thus, the claimed sub band energy is the high-band energy.

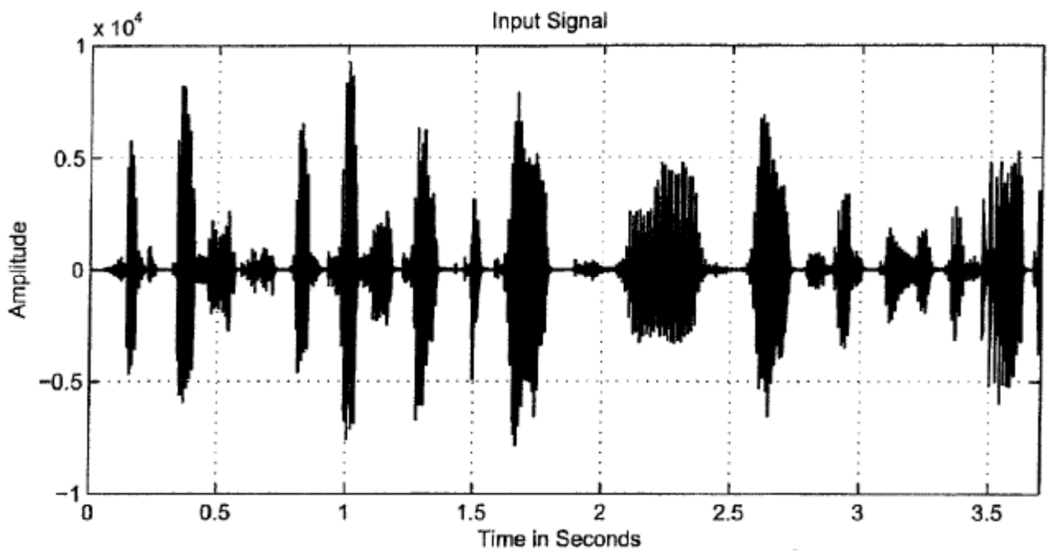
Nilsson discloses the high-band energy. For instance, *Nilsson* discloses that "[t]he **high-band energy y_0** is determined by computing a first LFCC using only a high-band part of the spectrum between fNu and fWu (where e.g. fNu=3,3 kHz and fWu=8,0 kHz)." Ex-1005*,* 11:44-55. Also, as discussed above, *Nilsson* disclose that Mel Frequency Spectral Coefficients (MFCC) is used as an alternative of LFCC. *Id.*, 8:45-50. Therefore, *Nilsson* inherently discloses computing Mel band energy level. Ex-1003, ¶118.

ii. *Iser* discloses that the sub band energy level value is attenuated when the power of the audio

signal approaches an estimate of the level of noise in the audio signal.

In the event that Patent Owner argues that *Nilsson* does not explicitly disclose that "the sub band energy level value is attenuated when the power of the audio signal approaches an estimate of the level of noise in the audio signal," *Iser* discloses it. Ex-1003, ¶119.

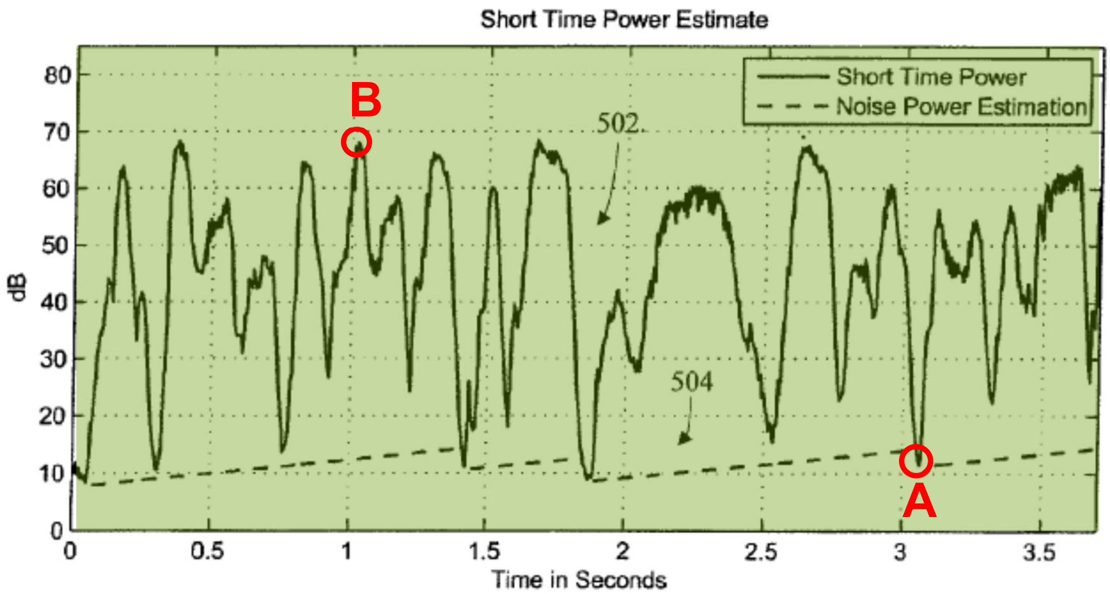
Iser discloses estimating power of the input audio signal and the noise power, and based on the estimation, applying "weighting factors to the signal component." For example, FIG. 4 shown below is "a representation of an input speech signal, such as the received acoustic signal x(n)." Ex-1006, 7:8-10.



Iser, FIG. 4

As shown below in annotated FIG. 5, *Iser* performs "a short time power estimation and a noise power estimation that correspond to the received acoustic signal x(n) of FIG. 4. Line 502 in FIG. 5 represents the estimated short time power

x(n) of the received acoustic signal x(n). Line 504 in FIG. 5 represents the noise power estimation b(n). The short time power estimation may be used to determine different factors for weighting the signal components." *Id.,* 7:8-18.



Iser, FIG. 5 (annotated)

Based on the estimated signal power and noise power, *Iser* further estimates "signal-to-noise ratio of the received acoustic signal". As shown above in annotated FIG. 5, when the power of the audio signal approaches noise power estimation, the low signal-to-noise ratio is low, for example, at point A. On the other hand, when the power of the audio signal is far away from noise power estimation, the signal-to-noise ratio is high, for example, at point B.

Iser further discloses that a "first weighting factor gsnr(n) that may be applied is a function of an estimated signal-to-noise ratio of the received acoustic signal x(n)" and "[i]f the received acoustic signal x(n) contains speech passages

with a **low signal-to-noise ratio**, then this weighting factor may be used to **damp** the upper bandwidth extension signal yhigh(n).” *Id.,* 7:20-32.

If the power of the audio signal approaches an estimate of the level of noise in the audio signal, the signal-to-noise ratio is low, and the upper bandwidth extension signal yhigh(n) is dampled. Therefore, *Iser* discloses that the sub band energy level value is attenuated when the power of the audio signal approaches an estimate of the level of noise in the audio signal. Ex-1003, ¶124.

iii. POSITA Would Have Combined *Nilsson* and *Iser*

As discussed above, both *Nilsson* and *Iser* are directed to a method for bandwidth extension to generate bandwidth extended signal using a narrow-band audio signal. Also, both *Nilsson* and *Iser* disclose adjusting the generated high band signal components.

While *Nilsson* discloses a broad concept of adjusting the high band energy, *Iser* discloses a specific method for adjusting, that is, by estimating signal power and noise power in the audio signal and determining a signal-to-noise ratio based on the power estimation and applying a weight factor corresponding to the signal-to-noise ratio.

Therefore, POSITA would have been motivated to apply the teaching of *Iser* to *Nilsson*’s system to estimate signal power and noise power, to determine signal-to-noise ratio, and to damp the sub band energy level value of *Nilsson* so that the

sub band energy level value is attenuated when the power of the audio signal approaches an estimate of the level of noise in the audio signal, as taught by *Iser*. Ex-1003, ¶126. By doing so, *Nilsson*'s system may obtain "a more natural output signal," as described by *Iser*. Ex-1006, 8:6-7, 8:58-59. Also, by doing so, *Nilsson*'s system may achieve its purpose, i.e., to "accurately describe [] the particular wide-band frequency component." Ex-1005, Claim 1. Ex-1003, ¶126.

C. Independent Claim 10

1. [10preamble] "An apparatus comprising at least one processor and at least one memory including computer code, the at least one memory and the computer code configured to with the at least one processor cause the apparatus to at least:"

Nilsson discloses an apparatus comprising at least one processor, as claimed, ("computer" of *Nilsson*) and at least one memory including computer code, as claimed, (*Nilsson*'s "memory of a computer, comprising software").

For instance, *Nilsson* discloses "the object is achieved by a computer program directly loadable into the internal **memory** of a **computer**, comprising **software** for performing the method described in the above paragraph when said program is run on a computer." Ex-1005, 3:25-29.

"The signal decoder receives a narrow-band acoustic signal aNB, either via a communication link (e.g. in PSTN) or from a **storage medium** (e.g. a digital memory)." *Id.*, 4:53-55.

2. [10a] "generate an excitation signal from an audio signal, wherein in the audio signal comprises a plurality of frequency components;"

As explained in Claim [1a], *Nilsson* discloses generating an excitation signal from an audio signal, wherein in the audio signal comprises a plurality of frequency components. *See* Section VII.B.2.

3. [10b] "extract a feature vector from the audio signal, wherein the feature vector comprises at least one frequency domain component feature and at least one time domain component feature;"

As explained in Claim [1b], *Nilsson* discloses extracting a feature vector from the audio signal, wherein the feature vector comprises at least one frequency domain component feature and at least one time domain component feature. *See* Section VII.B.3.

4. [10c] "determine at least one spectral shape parameter from the feature vector, wherein the at least one spectral shape parameter corresponds to a sub band signal comprising frequency components which belong to a further plurality of frequency components; and"

As explained in Claim [1c], *Nilsson* discloses determining at least one spectral shape parameter from the feature vector, wherein the at least one spectral shape parameter corresponds to a sub band signal comprising frequency components which belong to a further plurality of frequency components. *See* Section VII.B.4.

5. [10d] "generate the sub band signal by filtering the excitation signal through a filter bank and weighting the filtered excitation signal with the at least one spectral shape parameter,"

As explained in Claim [1d], *Nilsson* discloses generating the sub band signal by filtering the excitation signal through a filter bank and weighting the filtered excitation signal with the at least one spectral shape parameter. *See* Section VII.B.5.

6. [10e] "wherein the spectral shape parameter is a sub band energy level value and the sub band energy level value is attenuated when the power of the audio signal approaches an estimate of the level of noise in the audio signal."

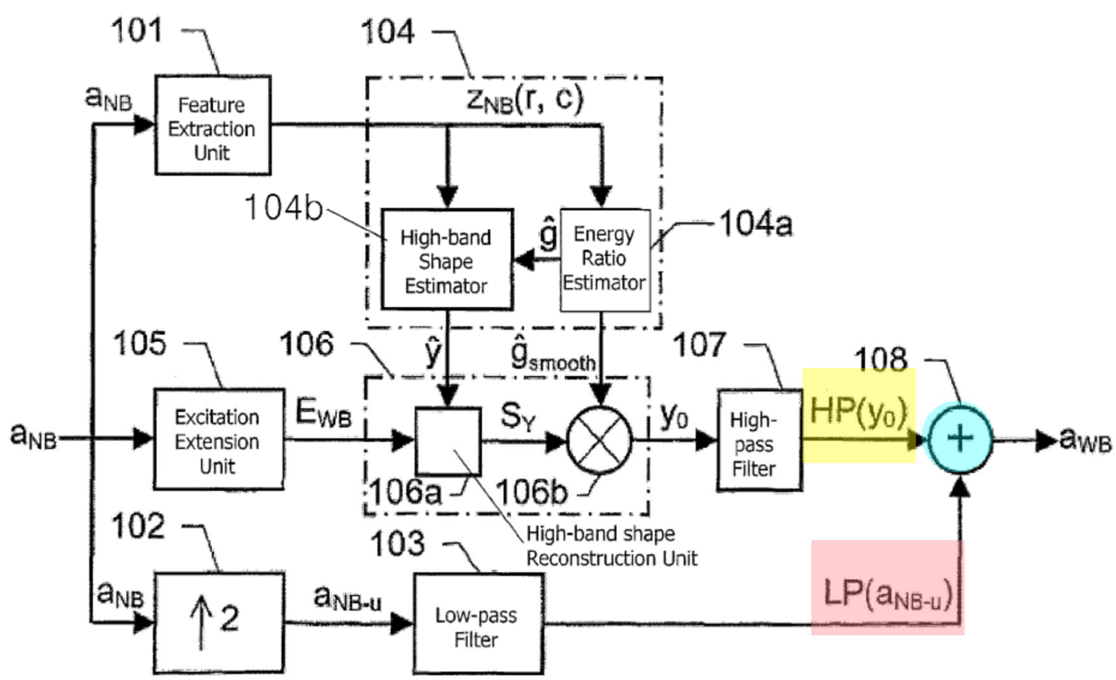
As explained in Claim [1e], the combination of Nilsson and *Iser* discloses that the spectral shape parameter is a sub band energy level value and the sub band energy level value is attenuated when the power of the audio signal approaches an estimate of the level of noise in the audio signal. *See* Section VII.B.6.

D. Dependent Claims 9 and 18

1. [Claim 9] "The method as claimed in any of claim 1, further comprising combining the sub band signal with the audio signal to provide a bandwidth extended audio signal."

Nilsson discloses combining the sub band signal with the audio signal to provide a bandwidth extended audio signal, as claimed, (*Nilsson*'s adder 108 combines the "low-pass filtered signal LP(a_{NB}-u)" and the "high-pass filtered signal HP(y0)" to form the "wide-band acoustic signal a_{WB}"). Ex-1003, ¶¶135-137.

As shown below in annotated FIG. 5, *Nilsson* discloses that "the adder 108 receives the low-pass filtered signal LP(a_{NB}-u), receives the high-pass filtered signal HP(y0) and adds the received signals together and thus forms the wide-band acoustic signal a_{WB}, which is delivered on the signal decoder's output." Ex-1005, 12:17-21.



Nilsson, FIG. 5 (annotated)

2. [Claim 18] "The apparatus as claimed in claim 10, wherein the at least one memory and the computer code is configured to: combine the sub band signal with the audio signal to provide a bandwidth extended audio signal."

As explained in Claim 9, *Nilsson* discloses combining the sub band signal with the audio signal to provide a bandwidth extended audio signal. *See* Section VII.D.1.

VIII. GROUND 2: *NILSSON*, *ISER*, AND *ETSI201.108* RENDER OBVIOUS CLAIMS 4, 5, 13, AND 14

1. POSITA Would Have Combined *Nilsson* with *Iser* and *ETSI201.108*

Both *Nilsson* and *Iser* are directed to a method for bandwidth extension to generate bandwidth extended audio signal using a narrow-band audio signal in a frequency range of 0-4 kHz. Ex-1003, ¶¶138-152. Each reference includes similar module with similar functions, for instance, both *Nilsson* and *Iser* disclose adjusting the high-band (e.g., 4-8 kHz) signal components. Ex-1003, ¶¶138-139. Thus, POSITA would have sought to combine *Nilsson* with *Iser* due to the compatible components. Ex-1003, ¶¶138-152.

Second, *Nilsson* discloses using Mel Frequency Spectral Coefficients (MFCC) to determine the spectral shape parameter for shaping the high-band signal. And, *ETSI201.108* discloses the well-known algorithms of computing MFCC. POSITA would have likewise sought standards like *ETSI201.108* for combining with *Nilsson* for these claimed limitations. Ex-1003, ¶139.

2. [Claim 4] "The method as claimed in claim 1, wherein the frequency components of the sub band signal are distributed according to a psychoacoustic scale comprising a plurality of overlapping bands, and frequency characteristics of the filter bank correspond to a distribution of frequency components of the sub band signal."

The '060 patent fails to define the "psychoacoustic scale." It appears that the claimed "psychoacoustic scale" is the psychoacoustically derived "Mel scale". Ex-1001, 11:3-6. The '060 patent also fails to define the "frequency characteristics of the filter bank." Regardless of this indefiniteness, this claim is unpatentable over *Nilsson* in view of *Iser* and *ETSI201.108*. Ex-1003, ¶¶140-147.

Nilsson discloses that the frequency components of the sub band signal are distributed according to Mel scale, and *ETSI201.108* discloses the well-known features of the Mel scale, e.g., a plurality of overlapping bands, and frequency characteristics of the Mel filter. Ex-1003, ¶141.

i. Mel scale is a psychoacoustic scale.

It is well-known in the art that Mel scale was derived historically from hearing-based psychoacoustic experiments that gave a perceptual measure of pitch, and therefore, Mel scale is a psychoacoustic scale. Ex-1012, Ex-1013. Ex-1003, ¶142.

ii. *Nilsson* discloses that the frequency components of the sub band signal are distributed according to Mel scale.

As discussed above, *Nilsson* discloses using Mel scale in representing the narrow-band spectral envelop c. For instance, *Nilsson* discloses that "[o]ther parameters than LFCCs can be used as alternative representations of the narrow-band spectral envelope x. Line Spectral Frequencies (LSF), **Mel Frequency Spectral Coefficients (MFCC)**, and Linear Prediction Coefficients (LPC) constitute such alternatives." Ex-1005, 8:46-54. Ex-1003, ¶143.

iii. Mel scale includes a plurality of overlapping bands, and frequency characteristics of the Mel filter correspond to a distribution of frequency components of the signal.

As discussed above, *Nilsson* discloses using Mel Frequency Spectral Coefficients (MFCC) in Mel scale to determine the spectral shape parameter. Therefore, *Nilsson* discloses all the inherent feature of Mel scale.

In the event that Patent Owner argues that *Nilsson* does not explicitly disclose the detailed feature of Mel scale, *ETSI201.108* discloses it. For instance, *ETSI201.108* discloses that Mel frequency cepstral coefficients are derived by a process in which an input signal "is divided into **overlapping frames of N samples**." Ex-1009, page 9, line 24. Ex-1003, ¶145.

ETSI201.108 also discloses the overlapping feature of Mel filter, which corresponds to the overlapping feature of the frames. For instance, *ETSI201.108*

discloses that "[t]he useful frequency band lies between 64 Hz and half of the actual sampling frequency. This band is divided into 23 channels equidistant in mel frequency domain. Each channel has triangular-shaped frequency window. Consecutive channels are half-**overlapping**." Ex-1009, page 10, line 25-27. Therefore, in Mel scale, the frequency characteristics of the Mel filter correspond to a distribution of frequency components of the sub band signal. Ex-1003, ¶146.

Therefore, the combination of *Nilsson* and *ETSI201.108* discloses that the frequency components of the sub band signal are distributed according to a psychoacoustic scale comprising a plurality of overlapping bands, and frequency characteristics of the filter bank correspond to a distribution of frequency components of the sub band signal. Ex-1003, ¶147.

3. [Claim 5] "The method as claimed in claim 4, wherein the overlapping bands are distributed according to the mel scale, and wherein the sub band signal is masked using at least one of: a triangular masking function; and a trapezoidal masking function."

As discussed above, *Nilsson* discloses the "Mel Frequency Spectral Coefficients (MFCC)" in which overlapping bands are distributed according to the Mel scale. Therefore, *Nilsson* discloses all the inherent feature of Mel scale, including the claimed feature. Ex-1003, ¶¶148-150.

In addition, *ETSI201.108* discloses this well-known feature of Mel scale. For instance, *ETSI201.108* discloses that "Mel Frequency Spectral Coefficients

(MFCC)" are derived by Mel filtering using "**triangular-shaped frequency window.**" Ex-1009, page 10, line 25-27. Ex-1003, ¶149.

Therefore, the combination of *Nilsson* and *ETSI201.108* discloses claim 5. Ex-1003, ¶150.

4. [Claim 13] "The apparatus as claimed in claim 10, wherein the frequency components of the sub band signal are distributed according to a psychoacoustic scale comprising a plurality of overlapping bands, and frequency characteristics of the filter bank correspond to a distribution of frequency components of the sub band signal."

As explained in Claim 4, the combination of *Nilsson* and *ETSI201.108* discloses that the frequency components of the sub band signal are distributed according to a psychoacoustic scale comprising a plurality of overlapping bands, and frequency characteristics of the filter bank correspond to a distribution of frequency components of the sub band signal. *See* Section VIII.2; Ex-1003, ¶151.

5. [Claim 14] "The apparatus as claimed in claim 13, wherein the overlapping bands are distributed according to the mel scale, and wherein the sub band signal is masked using at least one of a triangular masking function; and a trapezoidal masking function."

As explained in Claim 5, the combination of *Nilsson* and *ETSI201.108* discloses that the overlapping bands are distributed according to the mel scale, and wherein the sub band signal is masked using at least one of a triangular masking function; and a trapezoidal masking function. *See* Section VIII.3; Ex-1003, ¶152.

IX. GROUND 3: *NILSSON*, *ISER*, AND *VOS* RENDER OBVIOUS CLAIMS 2, 3, 7, 11, 12, AND 16

1. POSITA Would Have Combined *Nilsson* with *Iser* and *Vos*

Nilsson, *Iser*, and *Vos* all are directed to a method for bandwidth extension to generate bandwidth extended audio signal using a narrow-band audio signal in a frequency range of 0-4 kHz. Ex-1003, ¶¶153-187. Each reference includes similar module with similar functions, for instance, *Nilsson*, *Iser*, and *Vos* disclose adjusting the high-band (e.g., 4-8 kHz) signal components. Thus, POSITA would have sought to combine *Nilsson* with *Iser* and *Vos* due to the compatible components. Ex-1003, ¶¶153-154.

Second, both *Nilsson* and *Vos* are directed to generating excitation signal, up sampling the excitation signal, filtering, weighting the excitation signal. *Vos* discloses many different types of filters for generating and processing the excitation signal. POSITA would have implemented these filters in *Nilsson*'s system due to the advantages provided by the filters. The advantages of these various filters are well-known in the art and also mentioned by *Vos*. Ex-1003, ¶154.

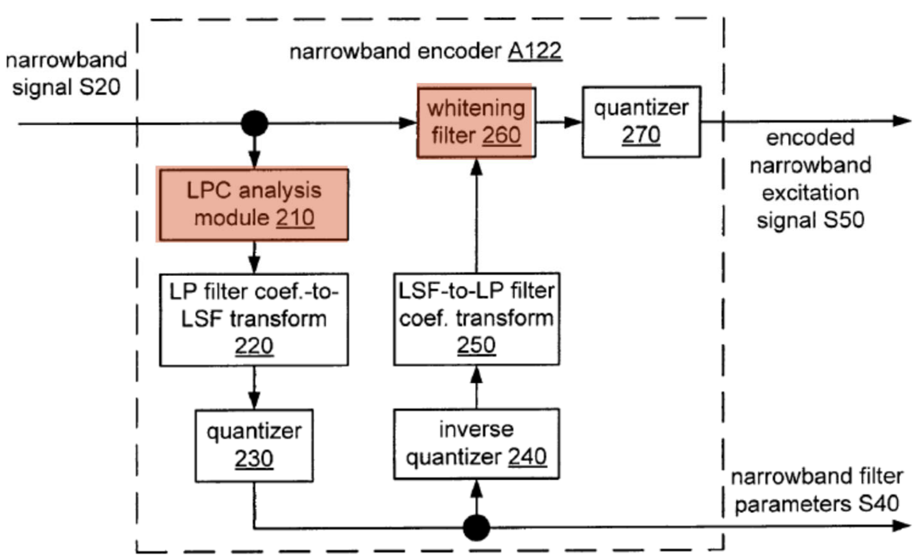
2. [Claim 2a] "The method as claimed in claim 1, wherein generating the excitation signal comprises: generating a residual signal by filtering the audio signal with an inverse linear predictive filter;"

The '060 patent discloses that "[o]nce the LPC filter coefficients have been determined within the excitation signal generator 417, the input audio signal frame

404 can in some embodiments be filtered by the LP analysis filter in order to produce a LP residual signal." Ex-1001, 22:54-57. Therefore, the inverse linear predictive filter is the LP analysis filter. Ex-1003, ¶¶155-163.

As discussed above, *Nilsson* discloses generating the excitation signal. In the event that Patent Owner argues that *Nilsson* does not explicitly disclose generating a residual signal by filtering the audio signal with an inverse linear predictive filter, then *Vos* further discloses this claim element. Ex-1003, ¶156.

Vos discloses generating a residual signal by filtering a narrowband signal with a whitening filter implemented with a linear prediction coding (LPC) analysis module, indicating the whitening filter is the inverse linear predictive filter. As shown below in annotated FIG. 6, *Vos* discloses a narrowband encoder A122 that generates encoded narrowband excitation signal S50 by filtering the narrowband signal 20 with a whitening filter 260.



Vos, FIG. 6a (annotated)

i. *Vos'* whitening filter generates a residual signal.

Vos discloses that the "narrowband encoder A122 generates a **residual signal** by passing narrowband signal S20 through a whitening filter 260 (also called an analysis or prediction error filter) that is configured according to the set of filter coefficients." Ex-1007, [0118].

ii. *Vos'* whitening filter is the inverse filter.

Vos discloses that the whitening filter 260 "**removes the spectral envelope** to spectrally flatten the signal. The resulting **whitened signal** (also called a **residual**) has less energy and thus less variance and is easier to encode than the original speech signal," and "[t]he peaks that characterize this spectral envelope represent resonances of the vocal tract and are called formants." *Id*., [0113]. Because whitening filter 260 generates the residual signal by removing the formants, the whitening filter 260 is the claimed **inverse filter**. Ex-1003, ¶160.

iii. *Vos'* whitening filter is the linear predictive filter.

As shown above in annotated FIG. 6a, *Vos* discloses that "a linear prediction coding (LPC) analysis module 210 encodes the spectral envelope of narrowband signal S20 as a set of **linear prediction (LP) coefficients** (e.g., coefficients of an all-pole filter 1/A(z))," and that "LPC analysis module 210 is configured to

calculate a set of ten LP filter coefficients to characterize the formant structure of each 20-millisecond frame." Ex-1007, [0114].

In particular, as shown above in annotated FIG. 6a the linear prediction filter coefficients are transformed into a set of corresponding "line spectral frequencies (LSFs)" by LP filter coef.-to-LSF transform 220, quantized by "quantizer 230", dequantized by "inverse quantizer 240" transformed to corresponding set of LP filter coefficients by LSF-to-LP filter coefficient transform 250, and "this set of coefficients is used to configure whitening filter 260 to generate the **residual signal** that is quantized by quantizer 270," to form encoded narrowband excitation signal. *Id.*, [0116]-[0119].

Therefore, *Vos* discloses generating a residual signal by filtering the audio signal with the inverse linear predictive filter.

3. [Claim 2b] "filtering the residual signal with a post filter stage comprising an auto regressive moving average filter based on the linear predictive filter; and"

The '060 patent fails to define the "post filter stage."

Nevertheless, *Vos* discloses filtering the residual signal with a post filter stage comprising an auto regressive moving average filter, as claimed, (syntheses filter A220 in FIG. 10 of *Vos*) based on the linear predictive filter. Ex-1003, ¶¶164-169.

As discussed above, the residual signal generated by the whitening filter 260 is quantized by quantizer 270 to form encoded narrowband excitation signal. As shown below in annotated FIG. 10, the formed encoded narrowband excitation signal is fed to a highband excitation generator A300 included in a highband encoder A200 to derive a highband excitation signal S120. The highband excitation signal is then sent a synthesis filter A220. Ex-1007, [0132], [0134].

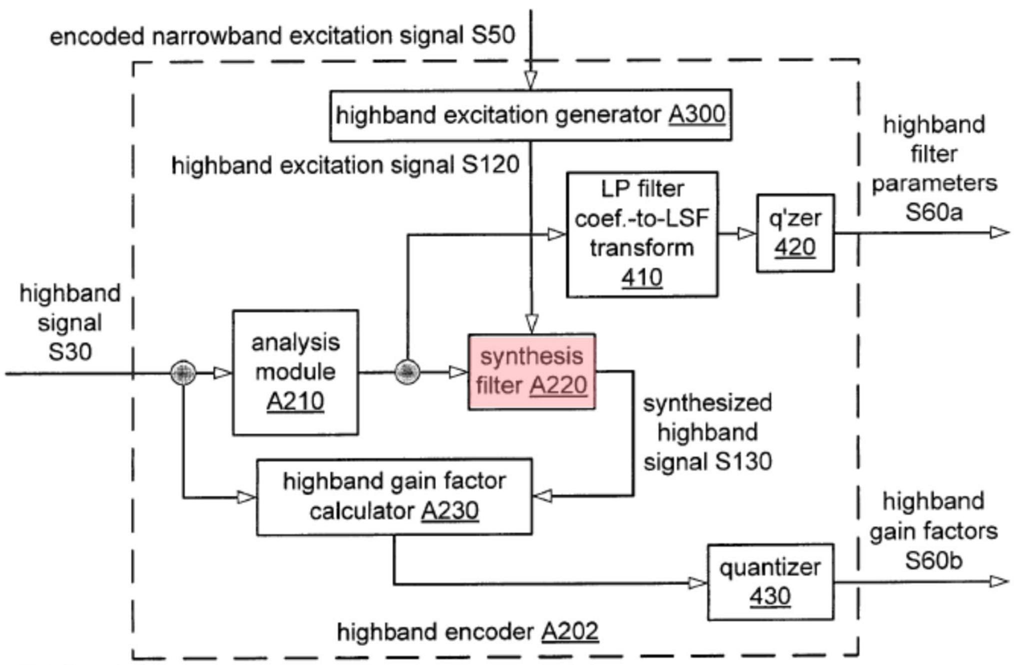


FIG. 10

Vos, FIG. 10 (annotated)

In particular, *Vos* discloses that the synthesis filter A220 is "configured to produce a synthesized highband signal S130 according to highband excitation signal S120 and the encoded spectral envelope (e.g., the set of LP filter coefficients) produced by analysis module A210. Synthesis filter A220 is typically

implemented as an IIR filter, although FIR implementations may also be used. In a particular example, synthesis filter A220 is implemented as a sixth-order **linear autoregressive filter**." *Id*., [0134].

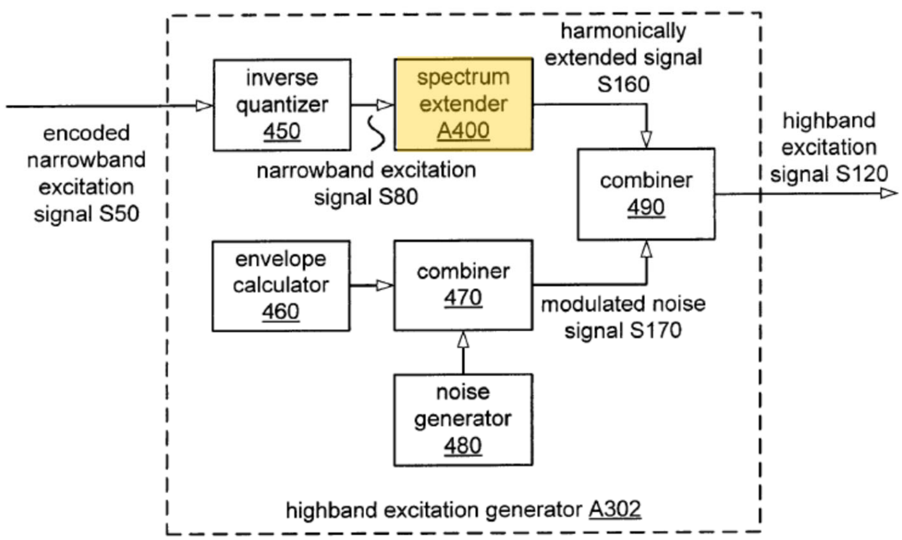
Therefore, the synthesis filter A220 of *Vos* is an autoregressive filter. Also, it is well known in the art that autoregressive model and moving average model are usually combined to form auto regressive moving average (ARMA) model and ARMA filter is well-known in the art. Ex-1014, Ex-1015. Ex-1003, ¶168. In this regard, ARMA filter can be said as a type of autoregressive filter. Because *Vos* discloses filtering the residual signal with auto regressive filter, POSITA would have understood that *Vos* also discloses filtering the residual signal with the auto regressive moving average filter. Ex-1003, ¶168.

Therefore, *Vos* discloses filtering the residual signal with a post filter stage comprising an auto regressive moving average filter based on the linear predictive filter. Ex-1003, ¶169.

4. [Claim 2c] "generating the excitation signal by up sampling and spectrally folding the output from the post filter stage."

Vos generates highband excitation signal by spectral folding operation or spectrally translating narrowband excitation signal S80 into the highband, as claimed, (e.g., via upsampling followed by multiplication with a constant-frequency cosine signal). Ex-1003, ¶¶170-173.

As shown above in annotated FIG. 10, the encoded narrowband excitation signal is fed to a highband excitation generator A300 included in a highband encoder A200 to derive a highband excitation signal S120. Ex-1007, [0132]. *Vos* further shows the details of the highband excitation generator A300 in FIG. 11. As shown below in annotated FIG. 11, the highband excitation generator A300 includes a spectrum extender A400 that "is configured to produce a harmonically **extended** signal S160 based on narrowband excitation signal S80." Ex-1007, [0139].



Vos, FIG. 11 (annotated)

For example, "spectrum extender A400 is configured to perform a **spectral folding operation** (also called mirroring) on narrowband excitation signal S80 to produce harmonically extended signal S160. Spectral folding may be performed by zero-stuffing excitation signal S80 and then applying a highpass filter to retain the

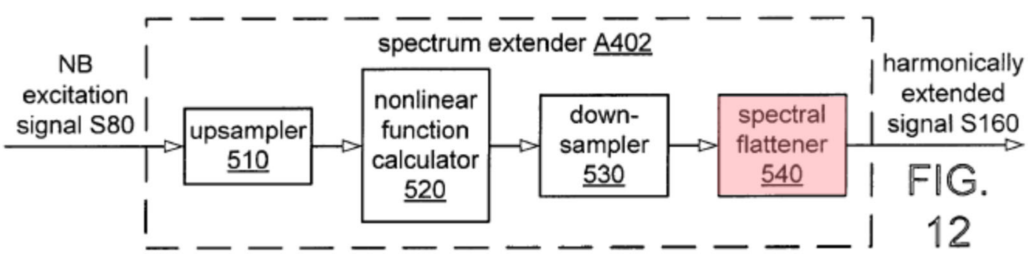
alias. In another example, spectrum extender A400 is configured to produce harmonically extended signal S160 by **spectrally translating** narrowband excitation signal S80 into the highband (e.g., via **upsampling** followed by multiplication with a constant-frequency cosine signal)." *Id.*, [0140].

Therefore, *Vos* discloses generating the excitation signal by up sampling and thus, spectrally folding the output from the post filter stage. Ex-1003, ¶173.

5. [Claim 3] "The method as claimed in claim 2, wherein the post filter stage further comprises a spectral tilt filter and a harmonic filter."

Vos discloses that the post filter stage further comprises a spectral tilt filter, as claimed, (spectral flattener in FIG. 12 of *Vos*) and a harmonic filter, as claimed, (the spectral flattener of *Vos*). Ex-1003, ¶¶174-177.

As shown below in annotated FIG. 12, *Vos* discloses a spectral flattener 540 included in spectral extender A402 and performs an adaptive whitening operation on the down-sampled signal.



Vos, FIG. 12 (annotated)

For example, *Vos* discloses that "spectral flattener 540 includes an LPC [linear prediction coding] analysis module configured to calculate a set of four

filter coefficients from the downsampled signal and a fourth-order analysis filter configured to whiten the signal according to those coefficients." Ex-1007, [0152].

Because the spectral flattener 540 performs whitening operation and flattens the spectral of the extended signal, the spectral flattener 540 is the spectral tilt filter. Also, because the spectral flattener 540 removes the harmonics, the spectral flattener 540 of *Vos* is the harmonic filter. Ex-1003, ¶177.

6. [Claim 7] "The method as claimed in claim 1, wherein the at least one frequency domain component feature of the feature vector comprises at least one of the following: a group of a plurality of energy levels of the audio signal, wherein each of the plurality energy levels corresponds to the energy of an overlapping band of the audio signal; a value representing a centroid of the frequency domain spectrum of the audio signal; and a value representing the degree of flatness of the frequency domain spectrum."

Nilsson and *Vos* disclose the elements of claim 7. Ex-1003, ¶¶178-183.

i. Energy levels

Nilsson discloses determining a group of a plurality of energy levels of the audio signal, wherein each of the plurality energy levels corresponds to the energy of an overlapping band of the audio signal.

As discussed above, *Nilsson* discloses a plurality of overlapping segments of an audio signal. Also as discussed above, *Nilsson* discloses determining c which is a vector of linear frequency cepstral coefficients. *Nilsson* further discloses that "[a]

first component c0 of the vector c constitutes the log **energy** of the narrow-band acoustic segment s." Ex-1005, 7:7-9.

Because *Nilsson* discloses calculating log energy of narrow-band acoustic segments, and the adjacent segments overlap, *Nilsson* discloses determining a group of a plurality of energy levels of the audio signal, wherein each of the plurality energy levels corresponds to the energy of an overlapping band of the audio signal. Ex-1003, ¶181.

ii. Centroid

In the event that the Patent Owner argues that *Nilsson* does not disclose this claim element, *Vos* explicitly discloses it. For instance, *Vos* discloses that "[t]he quantization maps each input value to a value that represents the corresponding Voronoi region (typically, the **centroid**), shown here as a point." Ex-1007*,* [0255]. Ex-1003, ¶182.

iii. Degree of flatness

In the event that the Patent Owner argues that *Nilsson* does not disclose this claim element, *Vos* explicitly discloses it. For instance, *Vos* discloses that "narrowband encoder A120 produces values for **spectral tilt** and speech mode parameters for each frame. Spectral tilt relates to the shape of the spectral envelope over the passband and is typically represented by the quantized first reflection coefficient. For most voiced sounds, the spectral energy decreases with increasing

frequency, such that the first reflection coefficient is negative and may approach −1. Most unvoiced sounds have a spectrum that is either flat, such that the first reflection coefficient is close to zero, or has more energy at high frequencies, such that the first reflection coefficient is positive and may approach +1." Ex-1007, [0130]. Ex-1003, ¶183.

7. [Claim 11a] "The apparatus as claimed in claim 10, wherein the at least one memory and the computer code configured to with the at least one processor to cause the apparatus to at least perform generating the excitation signal is further configured to: generate a residual signal by filtering the audio signal with an inverse linear predictive filter;"

As explained in Claim [2a], *Vos* discloses generating a residual signal by filtering the audio signal with an inverse linear predictive filter. *See* Section IX.2; Ex-1003, ¶183.

8. [Claim 11b] "filter the residual signal with a post filter stage comprising an auto regressive moving average filter based on the linear predictive filter; and"

As explained in Claim [2b], *Vos* discloses filtering the residual signal with a post filter stage comprising an auto regressive moving average filter based on the linear predictive filter. *See* Section IX.3; Ex-1003, ¶¶184.

9. [Claim 11c] "generate the excitation signal by up sampling and spectrally folding the output from the post filter stage."

As explained in Claim [2c], *Vos* discloses generating the excitation signal by up sampling and spectrally folding the output from the post filter stage. *See* Section IX.4; Ex-1003, ¶¶185-186.

10. [Claim 12] "The apparatus as claimed in claim 11, wherein the post filter stage further comprises a spectral tilt filter and a harmonic filter."

As explained in Claim 3, Vos discloses that the post filter stage further comprises a spectral tilt filter and a harmonic filter. *See* Section IX.5; Ex-1003, ¶186.

11. [Claim 16] "The apparatus as claimed in claim 10, wherein the at least one frequency domain component feature of the feature vector comprises at least one of the following: a group of a plurality of energy levels of the audio signal, wherein each of the plurality energy levels corresponds to the energy of an overlapping band of the audio signal; a value representing a centroid of the frequency domain spectrum of the audio signal; and; a value representing the degree of flatness of the frequency domain spectrum."

As explained in Claim 7, the combination of *Nilsson* and *Vos* discloses that the at least one frequency domain component feature of the feature vector comprises at least one of the following: a group of a plurality of energy levels of the audio signal, wherein each of the plurality energy levels corresponds to the energy of an overlapping band of the audio signal; a value representing a centroid of the frequency domain spectrum of the audio signal; and; a value representing the degree of flatness of the frequency domain spectrum. *See* Section IX.6; Ex-1003, ¶187.

X. GROUND 4: *NILSSON*, *ISER*, AND *KONTIO* RENDER OBVIOUS CLAIMS 6, 8, 15, AND 17

1. POSITA Would Have Combined *Nilsson* with *Iser* and *Kontio*

Nilsson, *Iser*, and *Kontio* all are directed to a method for bandwidth extension to generate bandwidth extended audio signal using a narrow-band audio signal in a frequency range of 0-4 kHz. Ex-1003, ¶¶188-206. Each reference includes similar module with similar functions, for instance, *Nilsson*, *Iser*, and *Kontio* disclose adjusting the high-band (e.g., 4-8 kHz) signal components. Thus, a POSITA would have sought to combine *Nilsson* with *Iser* and *Kontio* due to the compatible components. Ex-1003, ¶¶188-190.

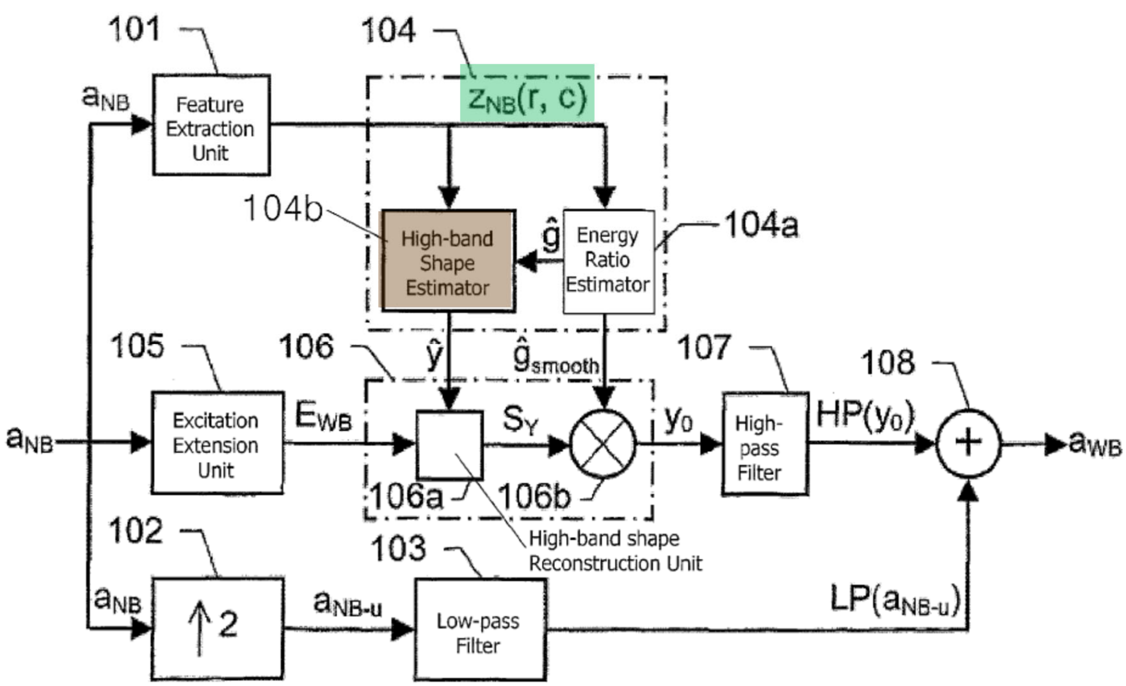
Second, both *Nilsson* and *Kontio* are directed to extracting feature vectors from the audio signal, determining spectral shape parameters from the feature vectors, and shaping the high-band signal using the spectral shape parameters. Ex-1003, ¶189.

In addition, *Kontio* discloses a neural network to provide the spectral shape parameters. *Kontio* also discloses extracting a gradient index. Considering the advantages provided by the neural network, POSITA would have adopted the neural network of *Kontio* in *Nilsson*'s system to produce the spectral shaper parameters. Also, considering the benefit provided by the gradient index, POSITA

would have extracted a gradient index in *Nilsson*'s system, as taught by *Kontio*, as discussed in detail below. Ex-1003, ¶190.

2. [Claim 6] "The method as claimed in claim 1, wherein determining at least one spectral shape parameter from the feature vector comprises: using a neural network to determine the at least one spectral shape from the feature vector, wherein the feature vector extracted from the audio signal forms an input target vector to the neural network, and wherein the neural network is trained to provide a sub band spectral shape parameter for the input target vector."

As shown below in annotated FIG. 5, *Nilsson* discloses determining the at least one spectral shape parameter $\hat{y}$ from the feature vector $z_{NB}(r, c)$, wherein the feature vector extracted from the audio signal forms an input target vector to a high-band shape estimator 104b.

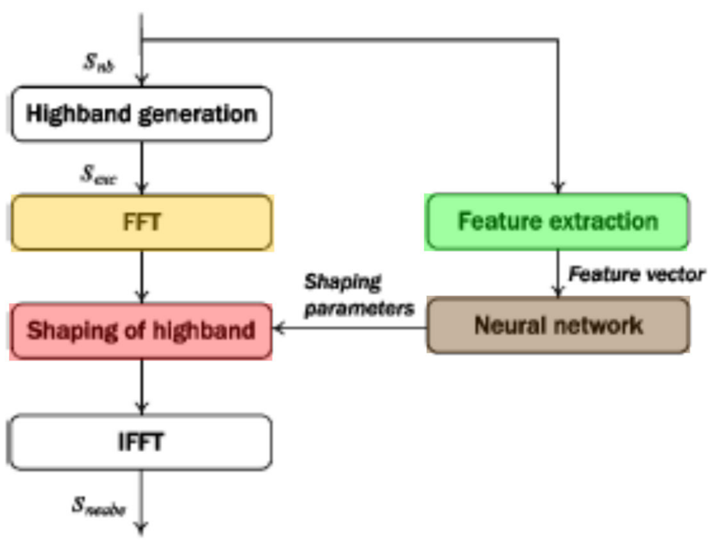


Nilsson, FIG. 5 (annotated)

In the event that Patent Owner argues that *Nilsson* does not explicitly disclose that the high-band shape estimator 104 is a neural network, *Kontio* explicitly discloses it. Ex-1003, ¶¶191-197.

i. ***Kontio* discloses using a neural network to determine the at least one spectral shape from the feature vector.**

As shown below in annotated FIG. 1, *Kontio* discloses using a neural network to determine spectral shape parameters. For instance, *Kontio* discloses that "[t]he **neural network** is the most crucial part of the algorithm, since it predicts the shape of the highband envelope, which, in turn, has a major impact on perceptual quality of the bandwidth-expanded sound." Ex-1008, page 874, 2:26-30. Because the high-band envelope is across all the high-band bandwidth, therefore, contains at least one spectral shape parameters.



Kontio, FIG. 1 (annotated)

Therefore, *Kontio* discloses using a neural network to determine the at least one spectral shape from the feature vector. In addition, using a neural network to determine spectral shape from the feature vector is well-known in the art. *See* Ex-1016 (in which *Iser*-2 also uses a neural network to determine spectral shape parameters). Ex-1003, ¶194.

ii. *Kontio* discloses that the feature vector extracted from the audio signal forms an input target vector to the neural network.

As shown above in annotated FIG. 1, *Kontio* discloses that the feature vector extracted from feature extraction module forms an input target vector to the neural network.

iii. *Kontio* discloses that the neural network is trained to provide a sub band spectral shape parameter for the input target vector.

Kontio discloses that "[t]he input to the neural network is computed solely from the narrowband speech signals which are obtained in the **training** phase by processing the corresponding wideband sounds." Ex-1008, page 876, 2:13-15. Also, as shown above in annotated FIG. 1, the output of *Kontio*'s neural network is at least one spectral shape parameter. Therefore, *Kontio* discloses that the neural network is trained to provide a sub band spectral shape parameter for the input target vector.

Because *Kontio* shows good results of bandwidth extension by adopting the neural network, and also the advantages of using neural network to determine high-band spectral shape parameters are well-known in the art (e.g., enhanced accuracy), POSITA would have motivated to implement the neural network of *Kontio* in high-band shaper estimator 104 of *Nilsson*. Ex-1003, ¶197.

Therefore, the combination of *Nilsson, Iser,* and *Kontio* render obvious claim 6. Ex-1003, ¶198.

3. [Claim 8] "The method as claimed in claim 1, wherein the at least one time domain component feature of the feature vector comprises at least one of the following: a gradient index based on the sum of the gradient at points in the audio signal which result in a change in direction of the waveform of the audio signal; a ratio of the energy of a frame of the audio signal to the energy of a previous frame of the audio signal; and; and a voice activity detector indicating whether a frame of the audio signal is classified as active or inactive."

Nilsson and *Kontio* disclose the elements of claim 8. Ex-1003, ¶¶199-204.

i. Gradient Index

In the event that the Patent Owner argues that *Nilsson* does not disclose this claim element directly, *Kontio* explicitly discloses it.

For instance, *Kontio* discloses that "To differentiate between voiced and unvoiced frames, a **gradient index** feature is used." Ex-1008, Page 875, 2:35-38.

ii. Energy ratio

Nilsson discloses the energy ratio. For instance, *Nilsson* discloses that "[t]he energy-ratio estimator 104a, which is included in the wide-band envelope estimator 104, receives the first component c0 in the vector of linear frequency cepstral coefficients c and produces, on basis thereof, plus on basis of the narrow-band shape x and the degree of voicing r an estimated **energy-ratio ĝ** between the high-band and the narrow-band." Ex-1005, 7:15-27.

In addition, *Kontio* also discloses the energy ratio. For instance, *Kontio* discloses that "frame energy ratio, takes into account temporal changes in the frame energy and is defined as the logarithmic ratio of the energies of two consecutive frames." Ex-1008, page 875, 2:6-8.

iii. Voice activity detector

Nilsson discloses that the feature extraction unit 101 determines "[t]he **degree of voicing r**, which represents one such essential feature zNB(r, c), [ ] by localising a maximum of a normalised autocorrelation function within a lag range corresponding to 50-400 Hz." Ex-1005, 6:25-32.

4. [Claim 15] "The apparatus as claimed in claim 10, wherein the at least one memory and the computer code configured to with the at least one processor to cause the apparatus to at least perform determining at least one spectral shape parameter from the feature vector is further configured to: use a neural network to determine the at least one spectral shape from the feature vector, wherein the feature vector extracted from the audio signal forms an input target vector to the neural network, and wherein the neural network is trained to provide a sub band spectral shape parameter for the input target vector."

As explained in Claim 6, the combination of *Nilsson* and *Kontio* discloses using a neural network to determine the at least one spectral shape from the feature vector, wherein the feature vector extracted from the audio signal forms an input target vector to the neural network, and wherein the neural network is trained to provide a sub band spectral shape parameter for the input target vector. *See* Section X.2; Ex-1003 ¶205.

5. [Claim 17] "The apparatus as claimed in claim 10, wherein the at least one time domain component feature of the feature vector comprises at least one of the following: a gradient index based on the sum of the gradient at points in the audio signal which result in a change in direction of the waveform of the audio signal; a ratio of the energy of a frame of the audio signal to the energy of a previous frame of the audio signal; and; a voice activity detector indicating whether a frame of the audio signal is classified as active or inactive."

As explained in Claim 8, the combination of *Nilsson* and *Kontio* discloses that the at least one time domain component feature of the feature vector comprises at least one of the following: a gradient index based on the sum of the gradient at

points in the audio signal which result in a change in direction of the waveform of the audio signal; a ratio of the energy of a frame of the audio signal to the energy of a previous frame of the audio signal; and; a voice activity detector indicating whether a frame of the audio signal is classified as active or inactive. *See* Section X.3; Ex-1003 ¶206.

XI. NON-INSTITUTION UNDER 35 U.S.C. §§ 314 OR 325 WOULD BE IMPROPER

Non-institution under 35 U.S.C. §§ 314(a) or 325(d) would be improper. The existence of parallel district court proceedings should not prevent institution of this Petition. *Cf. NHK Spring Co. v. Intri-Plex Techs., Inc.*, IPR2018-00752, Paper 8, at 19-20 (PTAB Sept. 12, 2018); *see also* Litigation. No factor favors denial of institution because, at this time, there is no investment beyond the initial discoveries at the district court and this IPR petition. Moreover, the strong merits of this case favor institution.

A. Non-Institution Under 35 U.S.C. §314(a) Is Improper

First, non-institution under 35 U.S.C. §§ 314(a) would be improper. Under the factors articulated in *Apple Inc. v. Fintiv, Inc.*, IPR2020-00019, Paper 11 (PTAB Mar. 20, 2020 (precedential), non-institution in light of the litigations would be improper because Factors 2-4 and 6 of the *Fintiv* factors favor institution, and Factors 1 and 5 are neutral.

Factor 1 (district court stay) is neutral. While ZTE moved to stay on December 30, 2020, *see* Litigation, Dkt. 46, there remains no indication that the district court will grant or deny the motion to stay. *Int'l Bus. Machines Corp. v. Trusted Knight Corp.*, IPR2020-00323, Paper 15 at 9 (PTAB Jul. 10, 2020). As discussed below, the motion to stay is based on ZTE's pending Motion to Dismiss

for Improper venue under § 1400, which unlike venue for convenience under § 1404, is not discretionary. *See* Litigation, Dkt. 44.

Factor 2 (proximity to district court trial) favors institution. The pending district court case is not scheduled for trial until June 20, 2022, and this date is subject to delays. Therefore, the Board will likely issue a final written decision before the pending district case. ZTE has moved to dismiss based on improper venue, and further moved to stay until venue is set. These motions are currently pending, making it "unclear that the court in the related district court litigation will adhere to any [future] scheduled jury trial date." *Sand Revolution II, LLC v. Cont'l Intermodal Group-Trucking LLC*, IPR2019-01393, Paper 24, at 9 (PTAB June 16, 2020) (informative). Should the district court grant either of the pending improper venue motion or the motion to stay, these cases will be assigned new, later trial dates likely in a new forum. Additionally, it is further noted that Judge Albright is unable to maintain trials based on their originally scheduled dates and is delaying the trials. As one example, the *VLSI Tech. LLC v. Intel Corp.*, No. 1:19-cv-00254 (W.D. Tex.) trial this month was delayed four months from its original date in October 2020.

Factor 3 (investment in district court case) favors institution. The parties' and the court's investment in this case has been minimal. *Fintiv*, Paper 11, at 11. Although preliminary infringement and invalidity contentions have been served,

claim construction has only just begun, and the claim construction hearing is not scheduled to occur for another two months in May 2021. In addition, the parties have not conducted any substantive fact discovery, as fact discovery does not open until after the claim construction hearing. Finally, the district court has not addressed the substance of the ’060 Patent—Patent Owner did not move for a preliminary injunction, and Defendants did not move to dismiss Patent Owner’s action based on the substance of the ’060 Patent, such as a motion to dismiss based on § 101. Where, as here, “the district court has not issued orders related to the patent at issue in the petition, this fact weighs against exercising discretion to deny institution.” *Fintiv*, Paper 11, at 10.

Factor 4 (overlapping issues) favors institution. WSOU is a prolific filer of patent infringement lawsuits. Based on WSOU’s litigation activity, it is likely that WSOU will bring more suits against other parties based on the ‘060 patent. Resolving the invalidity questions raised herein would mitigate any concern of duplicative efforts in the future. Additionally, ZTE will stipulate that, if this IPR is instituted, it will not pursue the specific grounds identified in this Petition (Sections VII-X) before the district court. This stipulation mitigates any concern of duplicative efforts. *Sand Revolution*, Paper 24, at 11-12.

Factor 5 (whether petitioner is also the defendant in district court) is neutral.

Factor 6 (other circumstances) favors institution. As explained above, the challenged claims are unpatentable over *Hwang*, *Kim*, *Miklos,* and *TS25.301*—none of which were considered during prosecution. A determination of its validity by the Board here would still save resources in the associated district court, and any additional cases WSOU may bring. There is a significant public interest against "leaving bad patents enforceable," *Thryv, Inc v. Click-To-Call Techs., LP*, 140 S. Ct. 1367, 1374 (2020).

B. Non-Institution Under 35 U.S.C. § 325 Is Improper

Second, non-institution under § 325 would also be improper based on a weighing of the factors set forth in *Becton, Dickinson & Co. v. B. Braun Melsungen AG*, IPR2017-01586, Paper 8 (PTAB Dec. 15, 2017). The asserted combinations are materially different and not cumulative of the prior art involved during the examination of the challenged claims. During prosecution, the following references were applied by the examiner, *Kornagel* (Ex-1017), *Iser-3* (Ex-1018). Both references are directed to bandwidth extension of audio signal.

The prior art references presented in this Petition were never listed by the Patent Owner nor cited by the examiner. They were never discussed or applied by the examiner to reject any claims. The references, for example, *Nilsson*, *Iser*, and *Vos*, are materially different from and not cumulative of the references cited in the prosecution at least because they disclose elements that were missing from

Kornagel and *Iser-3*. There is thus little to no overlap between the current and prior arguments. *Becton*, Paper 8, 23. But *Nilsson*, *Iser*, *Vos*, *Kontio* and *ETSI201.108* teach the limitations of ’060 patent, making denial under § 325(d) improper. Sections VII-X.

XII. MANDATORY NOTICES UNDER 37 C.F.R. § 42.8

A. Real Party-in-Interest

The real parties-in-interest are ZTE Corporation, ZTE (USA), Inc., and ZTE (TX), Inc.

B. Related Matters

Patent Owner has asserted the ’060 patent in litigation against Petitioners in the Litigation, filed on June 3, 2020. *See also WSOU Investments, LLC v. ZTE Corporation et al.*, 6:20-cv-00238 (WSOU initially asserted the ’060 patent against Petitioners on March 27, 2020 and dismissed the case on June 3, 2020).

C. Lead and Back-Up Counsel and Service Information

Petitioners provide the following counsel and service information. Pursuant to 37 C.F.R. § 42.10(b), Powers of Attorney accompany this Petition. Petitioners consent to e-mail service at the e-mail addresses identified in the table below, as well as at ZTE-WSOU-IPRs@finnegan.com.

LEAD COUNSEL (ZTE)	BACK-UP COUNSEL (ZTE)
Lionel Lavenue (Reg. No. 46,859) lionel.lavenue@finnegan.com Finnegan, Henderson, Farabow, Garrett & Dunner, LLP	Cory Bell (Reg. No. 75,096) cory.bell@finnegan.com Finnegan, Henderson, Farabow, Garrett & Dunner, LLP

1875 Explorer Street, Suite 800 Reston, VA 20190-6023 Tel: (571) 203-2750 Fax: (551) 203-2777	Two Seaport Lane Boston, MA 02210-2001 Tel: (617) 646-1641 Fax: (617) 646-1666 Bradford C. Schulz (Reg. No. 75,006) Bradford.schulz@finnegan.com Finnegan, Henderson, Farabow, Garrett, & Dunner LLP 1875 Explorer Street, Suite 800 Reston, VA 20190-6023 Tel: 571-203-2739 Fax: 202-408-440 Mingji Jin (Reg. No. 69,674) mingji.jin@finnegan.com Finnegan, Henderson, Farabow, Garrett & Dunner, LLP 901 New York Avenue NW, Washington, DC 20001-4413 Tel: (202) 408-6020 Fax: (202) 408-4400 R. Maxwell Mauldin (Reg. No. 78,626) r.maxwell.mauldin@finnegan.com Finnegan, Henderson, Farabow, Garrett & Dunner, LLP 271 17th Street, NW, Suite 1400 Atlanta, GA 30363-6209 Tel: (404) 653-6400 Fax: (404) 653-6444

XIII. GROUNDS FOR STANDING

Petitioner certifies the ’060 patent is available for IPR and Petitioner is not barred or estopped from requesting IPR challenging the patent claims on the grounds identified in this Petition.

XIV. CONCLUSION

Petitioner has established a reasonable likelihood of prevailing with respect to each of the challenged claims 1-18 of the '060 patent. Petitioner therefore requests the Board institute *inter partes* review and cancel each of these claims as unpatentable.

The Office may charge any required fees for this proceeding to Deposit Account No. 06-0916.

Date: March 29, 2021

Respectfully Submitted,

/Lionel M. Lavenue/
Lionel M. Lavenue, Lead Counsel
Reg. No. 46,859

CERTIFICATION UNDER 37 C.F.R. § 42.24(d)

Pursuant to 37 C.F.R. § 42.24(a)(1)(i), the undersigned hereby certifies that the foregoing PETITION FOR *INTER PARTES* REVIEW contains 12,448 words, excluding the parts of this Petition that are exempted under 37 C.F.R. § 42.24(a), as measured by the word-processing system used to prepare this paper.

/Lionel M. Lavenue/
Lionel M. Lavenue, Lead Counsel
Reg. No. 46,859

CERTIFICATE OF SERVICE

The undersigned certifies that the foregoing Petition for *Inter Partes* Review, the associated Power of Attorney, and Exhibits 1001 through 1018 are being served on March 29, 2021, by Priority Mail Express or by means at least as fast as Priority Mail Express at the following address of record for the subject patent.

Guy Gosnell
Nokia Corporation and Alston & Bird LLP
c/o Alston & Bird LLP
One South at The Plaza, 101 South Tryon Street
Suite 4000
Charlotte NC 28280-4000

A courtesy copy has been served on litigation counsel at:

Ryan S. Loveless
Etheridge Law Group, PLLC
2600 E. Southlake Blvd., Suite 120 / 324
Southlake, TX 76092

/Daniel E. Doku/
Daniel E. Doku
Litigation Clerk

FINNEGAN, HENDERSON, FARABOW,
GARRETT & DUNNER, LLP